Seria *Psihologie* este coordonată de Adrian Neculau

VALENTIN CLOCOTICI este profesor la Facultatea de Informatică a Universităţii „Al.I. Cuza" din Iaşi, coautor al cărţii *Metode statistice aplicate în industria textilă*, Editura Tehnică, Bucureşti, 1988, precum şi al unor articole ştiinţifice privind analiza statistică a datelor din diferite domenii de cercetare (psihologie, sociologie, biologie, tehnică). Predă cursuri de introducere în analiza statistică a datelor şi de statistică multivariată. A participat la realizarea multor contracte de cercetare aplicată.

AUREL STAN este profesor la Facultatea de Psihologie şi Ştiinţe ale Educaţiei din cadrul Universităţii „Al.I. Cuza" din Iaşi, doctor în psihologie. Autor al unor lucrări ştiinţifice în domeniul psihosomaticii, resurselor umane, teoriei testului psihologic şi psihometriei, cu preocupări speciale în computerizarea examenului psihologic şi analiza datelor din domeniul investigaţiilor psihologice. Predă cursuri de statistică psihologică, teoria elaborării testului psihologic şi resurse umane. Pe lângă experienţa ştiinţifică, autorul a desfăşurat o activitate diversificată în practica selecţiei şi orientării profesionale.

© 2000 by Editura Polirom Iaşi

Editura POLIROM
Iaşi, B-dul Copou, nr. 4, P.O. BOX 266, 6600
Bucureşti, B-dul I.C. Brătianu nr. 6, et. 7

Descrierea CIP a Bibliotecii Naţionale:
CLOCOTICI, VALENTIN
Statistică aplicată în psihologie / Valentin Clocotici, Aurel Stan
Iaşi: Polirom, 2001
296 p.; 24 cm. – (Collegium. Psihologie)

ISBN: 973-683-577-4

I. Stan, Aurel

519.22:159.9

Printed in ROMANIA

Valentin Clocotici **Aurel Stan**

STATISTICĂ APLICATĂ ÎN PSIHOLOGIE

POLIROM
2001

CUPRINS

Cuvânt înainte 9

Capitolul 1. NOŢIUNI INTRODUCTIVE 11

I.1. Statistica în ştiinţele sociale. Scurt istoric 11

I.2. Ce este statistica ? 13

I.3. Variabilitate şi măsurare 15

Capitolul 2. SCALE DE MĂSURĂ ŞI STATISTICI 20

2.1. Scala nominală 20

2.1.1. Alcătuirea unei scale nominale 20

2.1.2. Operaţii asupra categoriilor unei scale nominale 25

2.1.3. Statistici la nivelul unei scale nominale 27

2.2. Scala ordinală 30

2.2.1. Alcătuirea unei scale ordinale 30

2.2.2. Prezentarea datelor ordinale 32

2.2.3. Operaţii asupra scalelor ordinale 37

2.2.4. Statistici la nivelul unei scale ordinale 37

2.3. Scala de intervale 43

2.3.1. Alcătuirea unei scale de intervale 43

2.3.2. Prezentarea datelor de interval 46

2.3.3. Statistici pentru date de interval 52

2.4. Scala de raport sau scala de proporţii 57

Capitolul 3. INDICATORI DE MĂSURĂ A ÎMPRĂŞTIERII 63

3.1. Amplitudinea 64

3.2. Quantile 65

3.2.1. Abaterea quartilă 66

3.2.2. Criteriul 1,5×IQR 66

3.2.3. Rezumarea prin cinci valori 67

3.3. Abaterea medie absolută 67

3.4. Dispersia 68

3.5. Abaterea standard 71

3.6. Reuniunea seriilor de date 72

3.7. Proprietăţi ale abaterii standard legate de curba normală 73

3.8. Construcţia baremelor bazate pe medie şi abatere standard 74

3.9. Transformări de variabile 77

3.9.1. Variabile normate 77

3.9.2. Transformări liniare 78

Capitolul 4. PREZENTAREA ŞI SISTEMATIZAREA DATELOR 83
4.1. Proceduri generale 83
4.2. Protocoalele de bază ale distribuţiilor 94
4.2.1. Protocolul de structură „subiect - rezultat total" 94
4.2.2. Protocolul de structură „subiect - grup de apartenenţă" 99
4.2.3. Protocolul de structură „subiect - condiţie de investigaţie" 101
Capitolul 5. NOŢIUNI DE TEORIA PROBABILITĂŢILOR 104
5.1. Modele teoretice 104
5.1.1. Raportul teorie - model 104
5.1.2. Modele în psihologie 105
5.2. Construcţii axiomatice 112
5.3. Evenimente 115
5.4. Probabilitate 118
5.5. Variabile aleatorii 120
5.5.1. Cazul finit 120
5.5.2. Cazul infinit 123
5.6. Valori tipice ale variabilelor aleatorii 125
5.6.1. Valoarea medie 125
5.6.2. Quantila de ordin α 126
5.6.3. Valoarea mod 126
5.6.4. Momentul de ordin r 127
5.6.5. Momentul centrat de ordin r 127
5.6.6. Dispersia 127
5.6.7. Abaterea standard 128
5.6.8. Coeficientul de variaţie 128
5.6.9. Coeficientul de asimetrie 129
5.6.10. Coeficientul de boltire. Excesul 130
5.7. Repartiţii teoretice 131
5.7.1. Repartiţia normală 131
5.7.2. Repartiţia log-normală 137
5.7.3. Repartiţia χ^2 137
5.7.4. Repartiţia Student (t) 138
5.7.5. Repartiţia Fisher-Snedecor (F) 140
5.7.6. Repartiţia binomială 141
5.7.7. Aplicaţie la repartizarea observaţiilor în clase 144
Capitolul 6. INFERENŢA STATISTICĂ. ESTIMAŢII 146
6.1. Populaţie statistică 146
6.2. Eşantion 148
6.3. Clasificarea populaţiei şi a eşantioanelor 149
6.4. Eşantionarea (sondajul) 151
6.4.1. Sondaje aleatorii 152
6.4.2. Sondaje nealeatorii 154
6.4.3. Erorile sistematice în sondaj 155
6.5. Acurateţe şi eroare în sondaj 156
6.6. Logica inferenţei statistice 157
6.7. Estimaţii 159
6.7.1. Estimaţii punctuale 160
6.7.2. Intervale de încredere 169
Capitolul 7. TESTAREA IPOTEZELOR STATISTICE 169
7.1. Ipoteze statistice 169
7.2. Teste statistice 170
7.2.1. Testul bilateral (A) 176
7.2.2. Testul unilateral (B) 177
7.2.3. Testul unilateral (C) 177
7.3. Erori 178
7.4. Testarea valorilor de sondaj aberante 180
7.4.1. Testul Dixon 181
7.4.2. Testul Grubbs 182
7.5. Teste de concordanţă 184
7.5.1. Testul erorii standard a mediei 185
7.5.2. Testul de concordanţă Student (t) 187
7.5.3. Testul erorii standard a unei proporţii 188
7.5.4. Testul Kolmogorov-Smirnov (K-S) 189
7.5.5. Testul de concordanţă χ^2 191
7.6. Teste de comparare 193
7.6.1. Situaţii de aplicare a testelor de comparare 193
7.6.2. Testul F 195
7.6.3. Testul Student pentru eşantioane independente, abateri standard egale 197
7.6.4. Testul Student pentru eşantioane independente, abateri standard inegale 198
7.6.5. Testul Z pentru eşantioane independente, dispersii cunoscute 199
7.6.6. Testul pentru două proporţii independente 199
7.6.7. Testul Student pentru eşantioane cu date perechi 203
7.6.8. Testul Z pentru două proporţii dependente 205
7.6.9. Testul Mann-Whitney (U) 207
7.6.10. Testul Kolmogorov-Smirnov (K-S) pentru două eşantioane 210
7.6.11. Testul Wilcoxon 213
7.6.12. Testul McNemar 217
7.6.13. Testul χ^2 de comparare (asociere) 218
7.6.14. Utilizarea testelor χ^2 în psihologie 224
7.7. Testarea normalităţii 227
7.7.1. Testul Lilliefors 228
7.7.2. Teste bazate pe coeficienţii beta 229
Capitolul 8. ASOCIEREA DATELOR 230
8.1. Generalităţi 230

8.2. Asocierea între două variabile nominale (calitative) 234
8.3. Asocierea datelor ordinale 242
8.3.1. Coeficientul de corelaţie a rangurilor (Spearman) r_s 242
8.3.2. Coeficientul gama (γ) 244
8.3.3. Coeficientul de concordanţă Kendall (K) 247
8.4. Asocierea datelor continue (de interval) 249
8.4.1. Dreapta de regresie 249
8.4.2. Coeficientul de corelaţie (liniară) 253
8.5. Corelaţie în psihologie - situaţii specifice 255
8.6. Corelaţii parţiale 257
Capitolul 9. ANALIZA DISPERSIONALĂ 259
9.1. Analiza dispersională unifactorială (ANOVA pe o cale) 261
9.1.1. Structura datelor 261
9.1.2. Modele ale analizei dispersionale unifactoriale 265
9.2. Analiza dispersională bifactorială (ANOVA pe două căi) 268
9.2.1. Structura datelor 269
9.2.2. Modele ale analizei dispersionale bifactoriale 274
9.3. Teste privind egalitatea dispersiilor unor populaţii normale 277
9.3.1. Testul Bartlett 277
9.3.2. Testele Hartley şi Cochran 278
9.4. Aplicaţie 280

Anexe 285
Bibliografie 293

CUVÂNT ÎNAINTE

Statistică aplicată în psihologie este gândită ca un text introductiv şi din această cauză accentul este pus pe prezentarea raţionamentului statistic şi a tehnicilor fundamentale de analiză a datelor experimentale.

Nevoia de cunoştinţe statistice în domeniul psihologiei aplicate a fost resimţită imediat după introducerea experimentului ca metodă de bază a investigaţiei psihologice, la sfârşitul secolului al XIX-lea şi începutul secolului XX. Actualmente nu există domeniu al cercetării psihologice sau al practicii profesionale care să nu apeleze la cunoştinţe de statistică şi analiză a datelor. Fără utilizarea adecvată a cunoştinţelor de statistică, rezultatele unei cercetări riscă să fie respinse de comunitatea ştiinţifică şi, simultan, poate conduce la receptarea simplistă a rezultatelor comunicate de alţi autori. În acest mod are loc o izolare de viaţa ştiinţifică în domeniul psihopedagogic şi nu numai. Aceleaşi cunoştinţe de statistică îi sunt strict necesare şi psihologului practician, care se confruntă în viaţa profesională cu nevoia stringentă de acumulare a datelor, de prelucrare, interpretare şi comunicare a rezultatelor, indiferent dacă lucrează în domeniul clinic, în consultanţa psihodiagnostică, în educaţie etc.

Datorită dezvoltării tehnicii de calcul şi extinderii utilizării acesteia în toate domeniile de activitate, în carte sunt prezentate cu precădere interpretările posibile ale rezultatelor statistice şi, nu în ultimul rând, modul cum se obţin aceste rezultate. Deşi nu este negat rolul matematicii în dezvoltarea tehnicilor statistice, nu trebuie să cădem în eroarea de „a identifica înţelegerea conceptuală cu înţelegerea matematică" (D.S. Moore, G.P. McCabe – *Introduction to the Practice of Statistics*). Din acest motiv parcurgerea acestei cărţii nu necesită cunoştinţe matematice avansate, formulele utilizate sunt date fără demonstraţii (disponibile în tratatele de statistică matematică), ceea ce facilitează înţelegerea aplicării tehnicilor statistice prezentate. Acest aspect este urmărit chiar în capitolul cel mai formalizat al lucrării, „Noţiuni de teoria probabilităţilor", în care se oferă doar suportul minim necesar unei înţelegeri a logicii raţionamentului statistic.

Deşi conceptele şi noţiunile statistice sunt introduse într-o ordine acceptată de comunitatea ştiinţifică, cele nouă capitole ale lucrării pot fi parcurse şi într-o ordine diferită, potrivit nevoilor cititorului. Astfel, parcurgerea materialului se poate opri la primele patru capitole pentru cititorul interesat doar de statistica descriptivă. Capitolele 8 – „Asocierea datelor" – şi 9 – „Analiza dispersională" – pot fi citite independent, necesitând doar cunoştinţe minimale privind testele statistice.

Capitolele 6 – „Inferenţa statistică. Estimaţii" – şi 7 – „Testarea ipotezelor statistice" – necesită totuşi o parcurgere prealabilă a capitolului 5 – „Noţiuni de teoria probabilităţilor" –, dar numai în măsura în care se doresc amănunte suplimentare privind noţiunile vehiculate. Trebuie totuşi menţionat faptul că testarea semnificaţiei statistice a rezultatelor

experimentale este, poate, domeniul cel mai important într-o analiză statistică şi de aceea testele statistice prezentate în capitolul 7 constituie nucleul „tare" al acestei lucrări.

Dat fiind scopul propus la elaborarea acestei lucrări, prezentarea tehnicilor de analiză multivariată a fost limitată doar la nivelul strict introductiv. Elaborarea unei lucrări dedicate tehnicilor de analiză multivariată a datelor figurează printre proiectele noastre de viitor.

Autorii

Capitolul 1

NOŢIUNI INTRODUCTIVE

Psihologia a fost acceptată abia în a doua jumătate a secolului al XIX-lea ca ştiinţă de sine stătătoare, o dată cu accentuarea laturii ei experimentale, fapt ce a deschis un câmp aplicativ din ce în ce mai extins. Începuturile au fost marcate de acumulări de observaţii şi date, deficitare sub aspectul interpretării lor, dacă le comparăm cu celelalte ştiinţe, care aveau un statut epistemologic clar delimitat şi acceptat de comunitatea ştiinţifică. Metoda observaţiei a fost şi rămâne una dintre cele mai uzitate metode de investigare în această ştiinţă, fapt care plasează psihologul, indubitabil, în domeniul relativului empiric, adică în lumea fenomenelor, a caracteristicilor şi a proprietăţilor, care sunt vag conturate şi neasamblate în sisteme de cunoştinţe riguros închegate.

Se perpetuează, chiar şi în zilele noastre, situaţia în care aserţiunile psihologului profesionist sunt bagatelizate, atunci când se apropie de modul comun de percepere, şi respinse ca profund neadevărate, atunci când contrazic structuri perceptive şi mentale comune.

Sistemul psihic uman este o realitate cel puţin la fel de complexă ca domeniile unor discipline pozitive de vârf, cum ar fi biologia celulară sau astronomia. Dacă aceste discipline utilizează din plin cele mai complexe metode de investigaţie ştiinţifică, este firesc ca psihologia să tindă spre crearea şi utilizarea unor concepte tot mai rafinate, în paralel cu dezvoltarea unei metodologii de investigaţii complexe şi personalizate, care păstrează continuu legătura cu substanţa *sui generis* a realităţii studiate.

Refuzul analizei cantitative în psihologie, sub motivaţia rigidizării unor fenomene de o deosebită variabilitate în manifestarea lor, nu se justifică în contextul cercetării psihologice contemporane.

Tehnicile statistice utilizate în cercetarea psihologică vin în întâmpinarea nevoii de diversificare a instrumentelor de explorare, conducând la o interpretare mai adecvată a rezultatelor investigaţiilor.

1.1. Statistica în ştiinţele sociale. Scurt istoric

Nu intenţionăm să facem o istorie completă a statisticii din domeniul psihosocial, dar considerăm că este util să punctăm câteva momente din evoluţia disciplinei respective. Precizăm că statistica psihologică a apărut ulterior celei din domeniul economic şi administrativ.

Preocupări relativ sistematice au fost prezente încă din secolul al XVII-lea, expresie a străduinței păturii negustorești de edificare a unui sistem de asigurări pe viață, prin crearea așa-numitelor „tabele de mortalitate" (J. Graunt, 1661). Prin intermediul acestora, un preot militar din armata lui Frederic cel Mare, J.P. Sußmilch, cercetează în 1741 „ordinea divină în distribuția sexului omenesc"[1].

Schimbul de scrisori dintre Pascal și Fermat, din 1654, constituie începutul fundamentării teoriei matematice a verosimilității. Această direcție continuă cu J. Bernoulli (1713), De Moivre (1733), Gauss (1809) și Laplace (1812). În lucrările ultimilor trei se pun bazele conceptului de distribuție normală (gaussiană), celebrul „clopot al lui Gauss".

A. Wagner, în 1864, căuta să concluzioneze asupra legităților „în acțiunea umană aparent arbitrară".

C. Wolf susținea, în anul 1732, ideea posibilității de a crea o branșă matematică a psihologiei, care să se numească psihometrie: „Aceste teoreme aparțin psihometriei, știință care este pe cale să se constituie și care se va ocupa de cunoașterea matematică a spiritului uman"[2].

Pentru prima dată, în raport cu nevoile cercetării în psihologie, distribuția gaussiană este folosită de astronomul și statisticianul belgian A. Quetelet (1796-1874) și de cercetătorul englez al eredității, Francis Galton (1882-1911), inițiatorul reputatei școli psihometrice engleze, școală considerată un reper în domeniu.

Precizarea cantitativă a relațiilor-raporturi (corelații) scoate în prim-plan pe A. Bravais (1846), Galton (1886), K. Pearson (1896 – coeficientul momentului produselor) și C.E. Spearman (1904 – coeficientul de rang); ultimul inițiază și analiza bifactorială a inteligenței.

Dezvoltarea mai nouă a statisticii de cercetare, beneficiară a îmbogățirii esențiale a posibilităților experimentale, datorează mult savantului englez R.A. Fisher (1923, 1935).

Interesul psihologiei față de statistică a fost susținut, în spațiul german, în domeniul psihofizicii de către Weber și Fechner, iar în zona anglo-saxonă de J.P. Guilford, C.I. Mosier, F.M. Lord, N.R. Novick și H. Gulliksen, prin teoria clasică a testului psihologic. Teoria probabilistică a testului psihologic datorează mult matematicianului danez G. Rasch.

Aplicațiile moderne ale statisticii în psihologie se realizează și în teoriile moderne ale răspunsurilor la itemi (W.J. van der Linden, Ronald K. Hambleton, Gerhard H. Fischer, Elisabeth Seliger, J.O. Ramsay, Robert Mokken), în teoria generalizabilității (L.J. Cronbach, G.C. Gleser, H. Nanda, N. Rajaratnam) și în variantele moderne ale analizei factoriale (J.P. Benzécri).

1. După Peter R. Hofstätter, *Psychologie*, Fischer Taschenbuch Verlag, Frankfurt-pe-Main, 1971.
2. J.J. Bernier, B. Pietrulewicz, *La psychométrie*, Gaëtan Morin (éd.), Montréal, Paris, Casablanca, 1997, p. 5.

1.2. Ce este statistica?

Statistica este un corp de metode utilizate pentru a colecta, a descrie și a analiza date numerice din observații sau investigații științifice. Statistica se concentrează, prin urmare, asupra reprezentărilor numerice ale diferitelor aspecte ale realității, adică asupra a ceea ce poate fi numărat, măsurat sau cuantificat.

Aria de folosire a metodelor statistice ocupă, practic, întregul spectru al științelor contemporane, inclusiv științele dominate de rigurozitate, legitate și exactitate.

În toate domeniile preocupărilor de cercetare științifică aplicativă, colectarea și prelucrarea datelor constituie o parte importantă, dar nu esențială, deoarece anterior acestui demers trebuie să fie stabilite ipotezele care se vor a fi confirmate sau infirmate. Deși apariția calculatoarelor electronice favorizează poziția de „detectiv" a cercetătorului, adică strângerea datelor fără nici o idee prealabilă, urmând ca prelucrarea ulterioară să evidențieze tendințe și regularități, această atitudine nu trebuie încurajată. „Legitățile" evidențiate printr-un asemenea procedeu prezintă o mare probabilitate de a se constitui în artefacte. Demersurile exploratoare, fără o ipoteză prealabilă, cedează teren în favoarea celor confirmatoare, care pleacă de la o asemenea ipoteză. Rezultatele obținute prin analize statistice își diminuează sever aportul științific dacă nu sunt completate de analize de conținut specifice domeniului de utilizare.

Metodele statistice pot fi diferențiate în metode descriptive și metode inferențiale, apărute cronologic în această ordine.

Statistica descriptivă include metodele de caracterizare a faptelor, a fenomenelor și a conjuncturilor. Această situație este exprimată în principal prin valori numerice însoțite sau nu de reprezentări grafice adecvate. Descrierea exhaustivă a unui fenomen este practic imposibilă. De exemplu, descrierea reacțiilor unei persoane la situații angoasante este strict limitată de metodele de investigare pe care le avem momentan la dispoziție, deoarece o persoană nu poate fi supusă succesiv la toate situațiile care generează stres (etica profesională limitează clar situațiile în care un cercetător poate produce stresul prin diferite tehnici experimentale).

Pentru ilustrarea principalilor termeni utilizați în statistică, să luăm în considerare situația unui cercetător psihosocial care investighează o colectivitate umană urmând criteriile:

- mediul de proveniență a persoanelor;
- statutul socioeconomic;
- sexul;
- vârsta.

În continuare, această colectivitate va fi denumită și **populație statistică**. În general, prin populație statistică se va înțelege totalitatea persoanelor, a faptelor etc., care constituie obiectul de interes al cercetării. Trebuie să se facă distincția dintre termenul de populație statistică și termenul de populație, în accepțiunea

curentă a cuvântului. Noţiunea de populaţie statistică va fi tratată pe larg într-un capitol separat.

Scopul cercetării este să producă o bună descriere prin stabilirea procentajelor de bărbaţi şi de femei, diferenţiate pe categorii de vârstă, a sferelor ocupaţionale diferenţiate pe vârstă şi sex etc.

Fiecare dintre criteriile de diferenţiere este o **variabilă**, adică o entitate care ia diferite valori pentru componenţii colectivităţii. Astfel, „mediul de proveniență" este o variabilă care poate lua trei valori: *urban*, *rural*, *suburban*; variabila „sex" poate lua două valori: *feminin*, *masculin*; „vârsta" poate lua valori de la 0 ani la peste 100 de ani; „statutul socioeconomic" este o variabilă ale cărei valori sunt deja înregistrate în pagini întregi ale nomenclatoarelor.

Variaţia valorilor unei caracteristici date este conceptul primar al descrierii statistice. Pentru date constante, fără modificări datorate unor surse cunoscute sau necunoscute, nu are sens o analiză statistică.

Mulţimea valorilor înregistrate pentru o caracteristică particulară constituie **distribuţia** variabilei respective. Fiecare variabilă are o distribuţie specifică pentru o populaţie dată. De exemplu, persoanele din colectivitatea studiată sunt distribuite după valorile: urban, rural, suburban, într-o formă specifică, ce este desemnată ca fiind distribuţia populaţiei după mediul de proveniență. Dacă există 35% persoane provenite din mediul urban, 50% persoane provenite din mediul rural şi 15% persoane provenite din mediul suburban, atunci 35%, 50% şi, respectiv, 15% reprezintă distribuţia variabilei mediu de proveniență. Este evident că altă colectivitate poate avea o altă distribuţie a acestei variabile. În mod similar, se va putea vorbi de distribuţia variabilei sex, statut socioeconomic etc.

În cazul în care colectivitatea studiată conţine, de exemplu, 100 000 de persoane, descrierea în termenii precizaţi anterior devine practic imposibilă, deoarece resursele de timp şi de persoane, necesare pentru investigaţie, sunt limitate şi, chiar dacă nu ar fi aşa, trecerea timpului antrenează modificări ale valorilor deja înregistrate, ceea ce conduce la o imagine aproximativă.

Se impune în acest caz recurgerea la metode oportune, care să permită obţinerea unei bune estimări a situaţiei de ansamblu prin extrapolarea situaţiei existente într-o parte a colectivităţii.

Statistica inferenţială grupează acele metode şi tehnici de estimare a caracteristicilor unei populaţii statistice din observaţii efectuate asupra unei submulţimi de componenţi ai populaţiei.

Revenind la exemplul anterior, este mai eficient să operăm o selecţie a 500 de persoane din colectivitatea studiată, să investigăm aceste persoane şi din informaţiile obţinute să deducem, să extrapolăm distribuţiile variabilelor: mediu de proveniență, sex etc., pentru întreaga populaţie statistică.

Submulţimea utilizată este denumită **eşantion.** Dacă eşantionul este corect selectat şi sunt utilizate proceduri adecvate de derivare a rezultatelor, vom obţine o descriere suficient de exactă a întregii populaţii statistice.

Se impune totuşi menţinerea unei diferenţieri între caracteristicile unui eşantion şi cele ale întregii populaţii statistice. Dacă în cadrul populaţiei vorbim de **parametri** (valori fixe existente la un moment dat – de exemplu, în colectivitatea studiată există 48,2%[3] bărbaţi), în cadrul eşantionului vorbim de **estimaţii** (sau **statistici**). Astfel, în eşantionul de 500 de persoane există, de exemplu, 48,1% bărbaţi. Dacă am selecta un alt eşantion, ar fi posibilă o valoare de 48,4% bărbaţi etc., iar valorile astfel obţinute ar oscila în jurul parametrului respectiv. Această situaţie justifică intuitiv termenul de **estimaţie** pentru valoarea calculată într-un eşantion, ca aproximare a valorii reale (a parametrului).

Atât în cercetările moderne, cât şi în practica actuală a instituţiilor (guvernamentale şi neguvernamentale, de prognoze, de asigurări, de sondaje etc.), tehnicile de inferenţă statistică au o utilizare preponderentă faţă de cele descriptive. Unele tehnici au devenit atât de rafinate încât erorile sunt substanţial diminuate, fapt care permite ca, pe baza studierii unor eşantioane de volum mic, să fie adoptate decizii de mare responsabilitate socială şi economică.

1.3. Variabilitate şi măsurare

Studiul conduitelor psihologice se referă, prin excelenţă, la un domeniu de mare variabilitate. Variabilitatea intraindividuală (aceeaşi persoană, plasată în aceeaşi situaţie, reacţionează diferit) şi cea interindividuală (persoane diferite, plasate în aceeaşi situaţie, reacţionează diferit) sunt în fapt adevăruri banale. Specialistul, omul de ştiinţă, este însă în măsură să diagnosticheze şi să prevadă, pe baza **explorării sistematice** şi a unor **explorări specifice**, situaţia exactă în care se află plasat un subiect sau un grup de subiecţi (variabilitate intragrup şi variabilitate intergrup)[4].

Având în vedere precizările anterioare, după care statistica are drept concept primar variabilitatea datelor, putem deduce că studiul caracteristicilor psihologice poate utiliza, cu beneficiu evident, metode şi tehnici statistice. Folosirea metodelor statistice necesită însă separarea surselor de variabilitate în două categorii: **surse previzibile şi surse imprevizibile.**

Datele pe care le prelucrează un psiholog provin din metode de investigaţie specifice: experiment, test, observaţie, chestionar, convorbire psihologică, anchetă psihologică etc. Din însăşi metoda de investigaţie utilizată într-un studiu concret, psihologul se aşteaptă ca o parte din variaţia observată şi cuantificată să se datoreze premiselor pe care le-a luat în calcul la stabilirea ipotezei de lucru. De exemplu, un studiu al capacităţii operaţionale a gândirii la elevii de 10, 12 şi 14 ani, evidenţiază niveluri diferite, după vârstă (în cazul utilizării unor metode verificate). Deci, în

3. Valoare necunoscută, de altfel. Dacă dispunem de această valoare, nu mai are sens căutarea ei.
4. O prezentare detaliată a acestui subiect se găseşte în capitolul consacrat analizei dispersionale.

acest studiu, vârsta este o sursă previzibilă de variaţie. Însă este evident că această sursă nu explică întreaga variaţie a scorurilor obţinute, existând posibilitatea intervenţiei unor multiple alte surse. Unele dintre acestea nu sunt urmărite în mod sistematic (de exemplu, mediul sociofamilial nu a fost luat în considerare), iar despre altele, practic, nu ştim nimic.

Rezumând, putem stabili că:

1) **variaţiile previzibile** sunt asociate la surse sistematice (care au fost clar stabilite în ipoteza de lucru);
2) **variaţiile imprevizibile** nu sunt asociate unor surse sistematice. Experimentatorul nu este pregătit (sau este incapabil datorită necunoaşterii lor) să le pună în evidenţă. Limitarea naturală a cunoaşterii omeneşti face să putem presupune permanent şi existenţa unui posibil ansamblu de surse imprevizibile.

De remarcat că sursele imprevizibile pot deveni previzibile ca urmare a extinderii ipotezelor de lucru sau a progresului cunoaşterii umane.

Cum ambele categorii de surse acţionează împreună, o problemă importantă este aceea a stabilirii diferenţiate a ponderii fiecăreia dintre ele. Când experimentatorul sau anchetatorul ştiinţific pleacă de la ipoteza că mai multe surse de variaţie sistematică au efecte asupra observaţiilor sale, el începe, în general, prin a verifica dacă fiecare dintre ele are efectiv un rol semnificativ. Experimentatorul va putea apoi să determine care sunt sursele de variaţie cele mai importante, acelea care contribuie cel mai mult la variaţiile observaţiilor. Această indicaţie va fi adesea foarte importantă pentru el, care urmăreşte explicarea fenomenului (a variaţiei notelor în test) şi, uneori, obţinerea de concluzii cu acţiune practică. Este necesară folosirea metodelor de analiză statistică pentru a evalua într-un mod mai sigur şi mai precis **greutatea unei surse de variaţie**. Această greutate este întotdeauna relativă, raportată la ansamblul celorlalte surse.

Noţiunea de greutate a unei surse este uşor de concretizat, dacă se ia în considerare situaţia unui examen cu mai multe probe. În totalul punctelor (care constituie aici observaţia finală), o probă cu coeficientul 3 va avea o greutate de trei ori mai mare decât o probă cu coeficientul 1. Dacă cineva încearcă să explice sau să prevadă rezultatul examenului, el va trebui să cunoască acei coeficienţi cu care probele individuale sunt ponderate pentru a produce nota finală.

Aceeaşi problemă de evaluare a greutăţii surselor de variaţie apare şi în ceea ce priveşte sursele fortuite (întâmplătoare). Dacă ponderea lor este mare în raport cu sursele sistematice, atunci va fi dificil să se evidenţieze şi să se recunoască semnificaţia acestora din urmă. Dispozitivul de observaţie va fi puţin sensibil la variaţiile sistematice care vor fi „înecate" în variaţiile fortuite.

Pentru a putea fi prelucrate, rezultatele oricărei cercetări trebuie să se prezinte sub formă cuantificată. Din moment ce rezultatul unei probe psihologice, al unui experiment psihologic etc. este exprimat printr-un număr, în continuare suntem obligaţi să respectăm regulile care domină lumea numerelor şi procedurile ştiinţelor care le tratează.

Cuantificarea datelor rezultate dintr-o cercetare are la bază un proces de **măsurare**. Prin măsurare vom înţelege procesul de determinare a diferitelor valori ale unei variabile. Măsurarea a fost definită ca „atribuirea unor numere la aspecte ale obiectelor sau evenimentelor potrivit unei reguli"[5].

Definiţia afirmă că măsurarea reprezintă orice set consistent de reguli pentru asocierea de numere la fenomene. Utilizarea unor asemenea reguli constituie o acţiune integrată în viaţa noastră zilnică: ceasul, ruleta, termometrul etc. Numerele care sunt citite pe scalele acestor instrumente dau sensul timpului, lungimii, temperaturii etc., chiar dacă nu toţi suntem familiarizaţi cu regulile care au stat la baza creării lor. Este de la sine înţeles că aceste reguli există şi sunt foarte exact definite.

În psihologie, posibilitatea măsurării directe este foarte strict limitată. Atribuirea numerelor în cercetările psihologice poate fi invocată în:

- timpul necesar pentru rezolvarea unei sarcini (excitaţie şi reacţie);
- cuantificarea unei aserţiuni (stabilirea prezenţei sau absenţei unei trăsături);
- indicaţii dobândite despre frecvenţa unui fenomen de un anumit fel;
- reacţia corectă la un stimul de un anumit gen.

Nota finală a unui test se compune din acumularea unor componente elementare (itemi) care, în unele cazuri, sunt grupate în subteste.

În orice domeniu al realităţii, măsurarea este posibilă doar dacă avem la dispoziţie instrumente de măsură adecvate. De exemplu: pentru lungimi – rigla, pentru timp – cronometrul etc. Procesul de măsurare, cel puţin în situaţiile curente, constă în citirea pe scalele instrumentelor a unor valori reprezentând numărul de unităţi fundamentale de măsură (sau subunităţi ale lor). Reţinem din acest proces importanţa existenţei unei scale de măsură. De asemenea, trebuie precizat faptul că instrumentul de măsură nu este neapărat un aparat fizic. De exemplu, în psihologie şi sociologie, testul şi chestionarul constituie instrumente de măsură frecvent acceptate şi utilizate. Nota rezultată în urma aplicării şi corectării unui test este similară unei gradaţii de pe scala unui aparat de măsură.

Este foarte important ca instrumentele şi scalele de măsură să fie acceptate pentru a putea înlesni, pe de o parte, comunicarea între specialişti şi, pe de altă parte, comunicarea între aceştia şi beneficiarii competenţei lor. Această largă acceptare este necesară şi pentru că, în definitiv, orice unitate de măsură constituie rezultatul unei convenţii.

> **Observaţie.** În continuare, folosirea noţiunii de scală de măsură implică automat existenţa unui dispozitiv suport (aparat de măsură, chestionar etc.), chiar dacă acesta nu este explicit indicat.

Pentru a-şi îndeplini menirea, o scală de măsură trebuie să fie **consistentă** şi **corectă**. O scală are consistenţă internă dacă produce rezultate (aproape) identice, atunci când este folosită în mod repetat pentru acelaşi obiect sau fenomen. Dacă

5. Smith S. Stevens, „Measurement", în Garry M. Maranell, *Scaling: A Sourcebook for Behavioral Scientist*, Aldine Publishing Company, Chicago, 1974.

pentru un cronometru cerința pare riguros satisfăcută, nu același lucru se poate spune despre testele psihologice, care satisfac diferențiat condiția de consistență. Impunerea în timp a unei probe psihologice este rezultanta mai multor condiții, dintre care una o reprezintă și consistența verificată în multiple aplicări.

O scală de măsură este corectă dacă produce informația pe care o așteptăm de la ea. De exemplu, termometrul poate fi folosit pentru măsurarea temperaturii dintr-o cameră, dar nu și pentru a măsura confortul camerei[6]. În ceea ce privește corectitudinea testului psihologic ca instrument de măsură, facem precizarea că există discipline metodologice care se ocupă în special de această problemă. Este foarte dificil să proiectăm, de exemplu, un test care să măsoare inteligența, o noțiune psihologică foarte discutată, deoarece o importanță deosebită o are ipoteza teoretică de la care pornim. Nu înseamnă însă că practica științifică nu a selectat o serie de probe larg apreciate, cum ar fi proba WISC, care măsoară nivelul de inteligență la copii.

În afară de consistență și de corectitudine, cerințe intrinsece ale unei scale de măsură, mai apar exigențe de raportare la obiectele și fenomenele măsurate. Din acest punct de vedere, o scală trebuie să fie **exhaustivă** și **mutual exclusivă**. Prima proprietate arată că scala poate măsura toate entitățile cărora le este destinată, iar a doua proprietate – că fiecare entitate primește, în urma măsurării, o singură valoare.

În concluzie, o scală este structurată în mod adecvat dacă gradațiile (categoriile) ei sunt suficiente pentru a epuiza întregul evantai de posibilități ale domeniului măsurat și, în același timp, fiecare entitate măsurată se încadrează în una și numai una dintre gradațiile scalei.

În teoria măsurării sunt cunoscute următoarele tipuri principale de scale de măsură:

- scala nominală;
- scala ordinală;
- scala de interval;
- scala de raport.

Procesul de creare a scalelor de măsură este unul continuu, adaptat mereu nevoilor practicii științifice. Astfel, au intrat în practica exprimării rezultatelor cercetării scalele de log-interval (o variantă a scalei de interval) și cele absolute (care acordă o importanță deosebită semnificației unităților de măsură).

În ordinea enumerării lor, se poate evidenția o creștere a complexității, fapt concretizat în mulțimea diferențiată de proprietăți ale rezultatelor măsurării.

Unii autori, din considerente teoretice, au exigențe sporite asupra atribuțiilor unei scale și nu consideră scala nominală o scală de măsură, datorită „sărăciei" semnificațiilor numerelor rezultate[7].

În alte cazuri, se procedează la o diversificare a tipurilor de scale prin luarea în considerare a unei scale dihotomice (caz particular al scalei nominale) ori a scalei

6. Norman R. Kurtz, *Introduction to social statistics*, McGraw-Hill Book Company, New York, 1983.
7. A.S. Reber, *The Penguin Dictionary of Psychology*, Penguin Books, Londra, 1985.

numite de Heinrich Wottawa[8] „scala metrică ordonată", care ar deține, în opinia autorului, o poziție intermediară între scala ordinală și cea de interval.

Prezentarea în detaliu a fiecărei scale și a tratamentelor statistice permise face obiectul capitolului următor.

8. În *Psychologische Methodenlehre*, Juventa Verlag, München, 1977.

Capitolul 2

SCALE DE MĂSURĂ ŞI STATISTICI

2.1. Scala nominală

Această scală este cunoscută şi sub denumirile de **scală calitativă, categorială** sau **de clasificare.** Este cel mai simplu tip de scală şi presupune doar diferenţierea calitativă a fenomenelor şi a obiectelor măsurate. Scala nominală constă în categorii care permit clasificarea (sortarea) obiectelor sau fenomenelor după o caracteristică sau un atribut.

2.1.1. Alcătuirea unei scale nominale

În cazul în care practicianul îşi propune să studieze populaţii caracterizate de variabile calitative, cum ar fi: categoria profesională, sexul, tipul bio-constituţional etc., el este nevoit să-şi alcătuiască pentru fiecare variabilă o scală de măsură care să-i permită înregistrarea observaţiilor. Scala adecvată o constituie, în acest caz, pentru variabilele calitative, cea nominală. Pentru a alcătui o asemenea scală, se procedează la inventarierea cât mai completă a întregului evantai de caracteristici posibile. Fiecare categorie distinctă de răspuns constituie o gradaţie a scalei. Mulţimea tuturor categoriilor astfel identificate formează scala nominală.

Un exemplu banal este scala nominală utilizată în stabilirea tipurilor temperamentale, aşa cum rezultă din clasificarea efectuată, încă din Antichitate, de Galenus şi Hypocrat:

- coleric;
- sangvinic;
- melancolic;
- flegmatic.

Deoarece fiecare individ poate fi „încadrat" într-unul dintre tipurile enumerate şi numai într-unul, cele patru tipuri pot constitui mulţimea categoriilor unei scale nominale, ce respectă cerinţele precizate pentru o scală de măsură. Facem precizarea că avem cunoştinţă de faptul că doar un procent relativ redus de persoane (aproximativ 15%) poate fi încadrat în tipurile „pure", majoritatea oamenilor fiind, de fapt, mixturi *sui generis* de tipuri temperamentale. Ca atare, încadrarea în categoria „coleric" înseamnă „mai mult coleric decât altceva" sau „predominant coleric".

Uneori, pentru ca o scală astfel alcătuită să îndeplinească condiţia de exhaustivitate, se introduce o categorie aparte, intitulată „alte situaţii", „şi alţii" etc. Pentru o cercetare eficientă, se impune ca această categorie specială să nu includă o parte semnificativă a populaţiei studiate, fapt ce ar diminua relevanţa rezultatelor.

În operaţiunea de culegere şi prelucrare a datelor se utilizează frecvent **codificarea datelor.** Prin codificare înţelegem transformarea unei mulţimi de obiecte, date, itemi, dintr-o formă sistematică de prezentare în altă formă sistematică de prezentare. În statistică, această codificare este necesară pentru a simplifica prelucrările datelor, atât prin procedeele clasice, cât şi prin cele electronice. De exemplu, tipurile temperamentale, enumerate anterior, pot fi codificate în felul următor:

Tabelul 2.1. *Exemplu de codificare*

Tip temperamental	Cod
coleric	1
sangvinic	2
melancolic	3
flegmatic	4

În locul codurilor 1, 2, 3, 4 putem utiliza a, b, c, d sau a_1, a_2, a_3, a_4 etc. Codificarea la scala nominală poate fi numerică, alfabetică, alfanumerică şi chiar pictografică (mai rar utilizată).

Codurile utilizate în scala nominală trebuie considerate simple simboluri, având unica proprietate de a fi diferite între ele, ceea ce permite discriminarea categoriilor.

Trebuie accentuat faptul că între categoriile („gradaţiile") unei scale nominale nu există relaţie de ierarhie, ordinea prezentării lor fiind indiferentă, lipsită de importanţă din punctul de vedere al consistenţei scalei. Utilizarea codificării poate crea iluzia unei ierarhii (1 este mai mic decât 2 etc., a se află înaintea lui b etc.), dar aceasta este doar o iluzie, utilizată uneori în scopuri de manipulare.

Exemplul 2.1. Să presupunem că ne propunem să caracterizăm 150 de indivizi după apartenenţa la un tip de temperament. După investigaţie, rezultă 32 de colerici, 48 de sangvinici, 37 de melancolici, 33 de flegmatici. Potrivit codificării de mai sus, datele pot fi prezentate astfel:

Tabelul 2.2. *Distribuţia tipurilor temperamentale (studiul 1)*

Tip temperamental	Număr de indivizi
1	32
2	48
3	37
4	33

Deşi avem două coloane de numere, menţionăm că în stânga avem coduri, etichete lipsite de semnificaţie cantitativă, în timp ce numerele din dreapta posedă o astfel de semnificaţie.

Ansamblul numerelor din coloana a doua poartă numele de **distribuţie** (sau **distribuţie de frecvenţe**) a tipurilor temperamentale în grupul studiat. Fiecare număr constituie **frecvenţa absolută** a diviziunii respective a scalei.

Precizăm încă o dată că ordinea categoriilor nu are importanţă în sine şi deci aceeaşi distribuţie poate fi dată astfel:

Tabelul 2.3. *Distribuţia tipurilor temperamentale (studiul 1a)*

Tip temperamental	Număr de indivizi
2	48
3	37
4	33
1	32

Să presupunem că un alt cercetător a obţinut pentru un grup de 250 de persoane următoarea distribuţie:

Tabelul 2.4. *Distribuţia tipurilor temperamentale (studiul 2)*

Tip temperamental	Număr de indivizi
1	58
2	73
3	80
4	39

O necesitate curentă a cercetării o constituie şi comparaţia cu unele cercetări similare. Compararea frecvenţelor absolute din cele două loturi (tabelele 2.2 şi 2.4) ne conduce la afirmaţia: „al doilea grup prezintă valori mai mari la toate categoriile". Este evident că „descoperirea" aceasta nu are valoare ştiinţifică, deoarece al doilea lot are un efectiv mai mare decât primul şi, prin urmare, este normal ca frecvenţele să fie mai mari. Pentru a fi posibilă o comparaţie eficientă, este necesar să utilizăm frecvenţele relative. **Frecvenţa relativă** pentru o categorie se obţine prin împărţirea frecvenţei absolute la numărul total de subiecţi. Prin înmulţirea cu 100, frecvenţele relative se exprimă ca **procentaje**.

Comparaţiile efectuate după frecvenţele relative (sau procentaje) capătă substanţă. Astfel, se poate afirma că proporţiile de colerici sunt aproximativ egale în cele două grupuri studiate (21,33%, respectiv 23,2%) sau că în grupul al doilea există mai mulţi melancolici decât în primul grup (32% faţă de 24,67%) etc.

Tabelul 2.5. *Distribuţia tipurilor temperamentale (comparaţie între cazul 1 şi cazul 2)*

	Grupul 1			Grupul 2		
Tip	Frecvenţa absolută	Frecvenţa relativă	Procent	Frecvenţa absolută	Frecvenţa relativă	Procent
1	32	32/150 =0,21	21,33%	58	58/250 =0,23	23,3%
2	48	48/150 =0,32	32,0%	73	73/250 =0,29	29,2%
3	37	37/150 =0,25	24,67%	80	80/250 =0,32	32,0%
4	33	33/150 =0,22	22,00%	39	39/250 =0,16	15,6%

Prin trecerea la frecvenţe relative (procentaje), măsurătorile efectuate în eşantioane de volume diferite sunt transformate într-un sistem de raportare comun, ceea ce permite comparaţii.

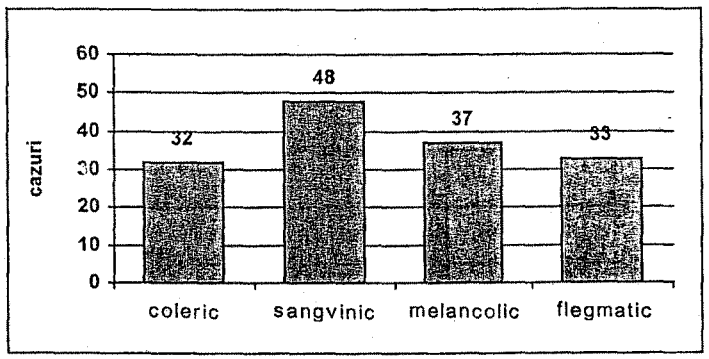


Figura 2.1. *Histograma frecvenţelor absolute*

Trebuie însă să ţinem seama de faptul că frecvenţele relative (procentajele) sunt sensibile la volumul eşantionului. Astfel, 5 cazuri reprezintă 50% din 10 cazuri, în timp ce 6 cazuri reprezintă 60% din 10 cazuri. Diferenţa de 1 între frecvenţele absolute se reflectă într-o diferenţă de 10 între procentaje. Dacă volumul eşantionului ar fi fost de 100, atunci 5 şi, respectiv, 6 cazuri ar fi reprezentat 5% şi, respectiv, 6%. Se observă diferenţele mai mici între procentaje, datorate numărului mai mare de cazuri din eşantion. Din acest motiv, nu se recomandă utilizarea procentajelor în raportări pentru loturi mai mici de 100 de observaţii (recomandare puţin respectată în cercetările curente). Această exigenţă are ca implicaţie practică cerinţa de a preciza întotdeauna numărul total de cazuri (volumul eşantionului), atunci când se raportează frecvenţe relative sau procentaje.

Distribuţiile pot fi vizualizate sub formă de reprezentări grafice specifice. Utilizarea unor asemenea reprezentări îmbogăţeşte sensibil transmiterea informaţiilor obţinute ca urmare a unor prelucrări de date.

Histograma este acea reprezentare grafică în care o serie de date este figurată, într-un sistem de referinţă bi- sau tridimensional, prin dreptunghiuri de înălţimi proporţionale cu valorile elementare ale seriei. În cazul general, axa X conţine gradaţiile scalei de măsură, iar axa Y gradaţiile corespunzătoare seriei de date.

Pentru o distribuţie obţinută după o scală nominală, gradaţiile axei X sunt categoriile scalei, fiecare reprezentată printr-un segment, iar axa Y este marcată cu frecvenţe absolute, frecvenţe relative sau procentaje. Potrivit numerelor de pe axa Y vom vorbi despre **histograma frecvenţelor absolute, histograma frecvenţelor relative** sau **histograma procentajelor**. În graficele din figurile 2.1 şi 2.2 sunt prezentate două exemple de histograme pentru datele exemplului 2.1.

Să remarcăm în aceste grafice faptul, specific scalelor nominale, că un dreptunghi nu are ca bază întregul segment de pe axa X, ci este situat în interiorul segmentului corespunzător. Aceasta pentru a ilustra caracterul discret (categorii distincte, fără elemente intermediare) al scalei nominale.

O a doua observaţie importantă este aceea că prezentările sub formă de diagramă cu bare verticale pot induce impresia falsă a unei ordini între categorii.

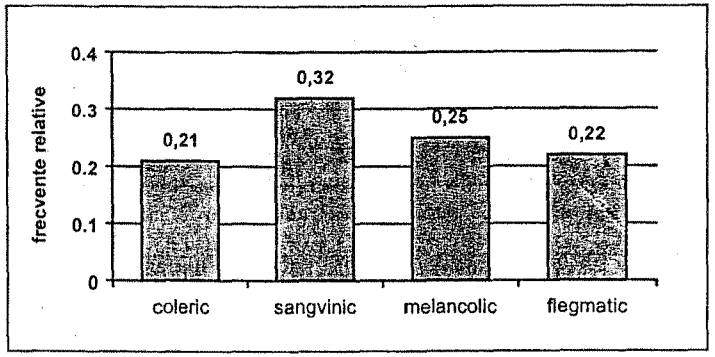


Figura 2.2. *Histograma frecvenţelor relative*

Din acest motiv, se recomandă în cazul scalelor nominale o aşezare răsturnată a histogramei, ca în figura 2.3.

Această histogramă, cunoscută sub numele de **diagramă cu bare**, reduce riscul sugerării unei ordini.

Cea mai adecvată formă de reprezentare grafică pentru datele unei scale nominale este diagrama circulară, de tip „clătită", diagramă „camembert", „pie" sau diagrama de structură, în care datele se reflectă în sectoare ale unui cerc (bidimensional) sau disc (tridimensional). În figura 2.4 se poate remarca aspectul *static*, *neordonat*, al unei asemenea reprezentări. Comparaţiile între numerele ilustrate sunt, uneori, mai greu de efectuat, dacă numerele respective sunt apropiate.

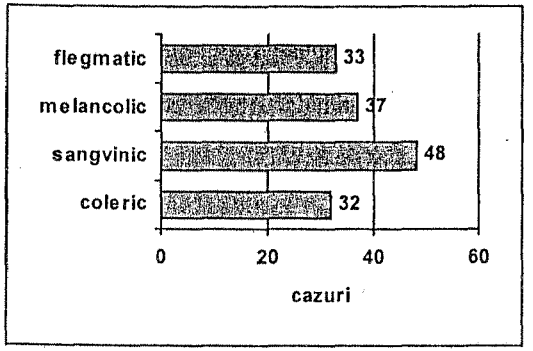


Figura 2.3. *Distribuţia tipurilor temperamentale (valori absolute)*

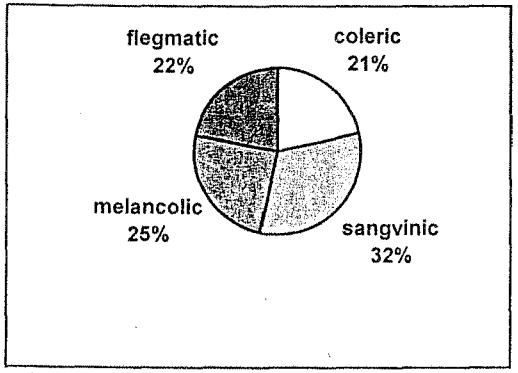


Figura 2.4. *Distribuţia tipurilor temperamentale (valori relative)*

Utilizarea reprezentărilor grafice în analiza datelor a consacrat, alături de alte argumente, noţiunea de **formă a unei distribuţii**. Precizăm că în cadrul scalelor nominale nu are sens noţiunea de formă a distribuţiei, datorită posibilităţii de interschimbare a localizării gradaţiilor (lipsa unei ordini între categorii este esenţială).

2.1.2. Operaţii asupra categoriilor unei scale nominale

Asupra categoriilor (diviziunilor) unei scale nominale se pot efectua două mari operaţii: gruparea (condensarea) şi rafinarea (diversificarea).

Prin **grupare** înţelegem acea operaţie prin care din două sau mai multe categorii se obţine o singură categorie, care va purta o etichetă nouă, distinctă de celelalte.

Prin repetarea operației de grupare, o scală nominală cu n categorii se va transforma într-o scală nominală cu m categorii, $m < n$.

Prin **rafinare** înțelegem operația prin care dintr-o categorie se obțin două sau mai multe categorii noi, distinct etichetate. Deci, prin repetarea operației, o scală nominală cu n categorii se transformă într-o scală nominală cu m categorii, $m > n$.

Asupra unei aceleiași scale nominale se pot efectua consecutiv atât operații de grupare, cât și operații de rafinare.

Operațiile prezentate corespund nevoilor practice de analiză și sinteză, generalizare și particularizare etc.

Toate aceste operații, abstract definite, nu pot fi totuși efectuate în mod arbitrar în cazul unei scale nominale concrete, responsabilitatea pentru întocmirea și/sau transformarea unei scale revenind cercetătorului avizat în domeniul psihologic, și nu statisticianului (chiar dacă aceștia reprezintă una și aceeași persoană), respectând exigențele de conținut și cele formal logice.

Ca exemplificare, prezentăm[9] categoriile de comunicare (Bales). Aceasta este o scală nominală, care permite clasarea comunicărilor dintr-un grup în funcție de rolul comunicării în procesul de discuție. Clasificarea furnizează informații despre dinamica proprie comunicării, raportul dintre participanții la dezbaterea în grup, mecanismele specifice discuției.

Setul extins de categorii, propus de Bales, este:

I.
1. manifestă solidaritate;
2. manifestă diminuarea tensiunii;
3. manifestă acordul, acceptă pasiv;
4. dă o sugestie;
5. emite o opinie;
6. dă o orientare, o informare;
7. cere o orientare, o informare;
8. cere o opinie;
9. cere o sugestie;
10. manifestă dezacordul;
11. exteriorizează tensiune;
12. manifestă antagonism.

Utilizând operațiunile de grupare, această scală se poate transforma, prin „condensarea" categoriilor, în modul următor:

II.
A. Reacții pozitive (din 1, 2 și 3)
B. Răspunsuri așteptate (din 4, 5 și 6)
C. Întrebări (din 7, 8 și 9)
D. Reacții negative (din 10, 11 și 12)

Observăm că noua scală cuprinde numai 4 categorii, fiecare provenind din câte trei categorii inițiale. O comunicare în cadrul unei dezbateri de grup poate fi măsurată (încadrată) atât prin intermediul scalei I, cât și al scalei II. Dacă, în primul

9. *Apud* M. Reuchlin, *Précis de statistique*, PUF, Paris, 1976, pp. 42-43.

caz, afirmația este încadrată în categoria 8 (cere o opinie), în scala II va fi încadrată în categoria C (întrebări). Folosirea scalei I sau a scalei II este impusă de nevoile pragmatice ale studiului.

Aceleași cerințe pot impune și un alt gen de grupare, cum ar fi:

III.
a. comunicare (6, 7)
b. evaluare (5, 8)
c. influență (4, 9)
d. decizie (3, 10)
e. tensiune (2, 11)
f. integrarea membrilor (1, 12)

Viziuni foarte sintetice ale fenomenului studiat pot produce condensări de ordinul al doilea (condensări ale unei scale deja condensate). Ca exemplu avem scala IV obținută din scala II:

IV.
i. reacții emoționale pozitive (A)
ii. reacții emoționale neutre (B, C)
iii. reacții emoționale negative (D)

Trecerea inversă, de la scale condensate la scale mai diversificate (cum ar fi de la scala II la scala I), se realizează prin operațiuni de rafinare și se impun din nevoi analitice accentuate.

Prin gruparea excesivă se va produce automat o pierdere de informații primare, în timp ce rafinarea excesivă va duce la o acumulare de amănunte care creează greutăți în perceperea esenței problemei studiate. De aici rezultă nevoia păstrării unui echilibru între cele două situații. Un element important în alegerea numărului de gradații ale unei scale nominale este și volumul eșantionului. Un eșantion relativ mic nu se poate studia folosind o scală nominală cu un număr mare de categorii; un eșantion mare va fi insuficient analizat pe o scală nominală cu puține categorii.

Efectul transformării unei scale asupra distribuției datelor unei cercetări concrete este marcant. Singurul lucru constant rămâne numărul total de membri ai eșantionului. Distribuirea acestora pe clase se va modifica la orice grupare sau rafinare a categoriilor. Reprezentările grafice vor urmări modificările de distribuție.

2.1.3. Statistici la nivelul unei scale nominale

Distribuțiile variabilelor urmărite în orice cercetare concretă constituie materialul primar cu care operează orice cercetător. De cele mai multe ori însă, eșantioanele luate în considerare sunt mari sau foarte mari, ceea ce duce la imposibilitatea tratării directe a informațiilor obținute. Apare astfel necesitatea „rezumării" acestor informații. În secțiunea 1.1.1. s-au prezentat asemenea rezumări ale unei distribuții atunci când s-au arătat reprezentările sub formă de tabel sau sub formă grafică a datelor rezultate în urma unei investigații. În secțiunea de față se vor introduce tehnicile care permit descrierea unei distribuții printr-o singură valoare numerică. O asemenea valoare, având rol de concentrare a informației, calculată pe baza datelor eșantionului, este numită **rezumat statistic** sau **statistică**. Nu trebuie omis faptul că cercetarea unui eșantion este efectuată în scopul obținerii de informații

asupra întregii populaţii statistice, ceea ce va conferi unui rezumat statistic caracterul de **estimaţie** a valorii adevărate care caracterizează populaţia statistică.

Scala nominală, fiind cea mai simplă în ierarhia după complexitate a scalelor de măsură, face posibile un număr restrâns de prelucrări şi rezumate statistice.

Valoarea **mod** (sau **modul, modul brut, moda**) este cea mai frecventă observaţie dintr-o distribuţie, altfel spus, categoria cu efectivul cel mai mare.

Exemplul 2.2. Să luăm în considerare că o cercetare asupra naturii intervenţiilor într-o dezbatere de grup a condus la obţinerea distribuţiei din tabelul 2.6, scala utilizată fiind scala II a lui Bales (vezi §2.1.2).

Tabelul 2.6. *Distribuţia intervenţiilor (scala II a lui Bales)*

Categoria de comunicare	Număr
A	73
B	55
C	40
D	62

Pentru această distribuţie, modul este categoria A care întruneşte cele mai frecvente intervenţii în discuţie, şi anume 73. Nu trebuie să ne surprindă similitudinea semantică a termenului **moda**, utilizat de unii autori pentru valoarea mod, cu cel din psihosociologia vieţii cotidiene.

O distribuţie cu o singură valoare mod se numeşte unimodală, iar una cu mai multe valori mod – multimodală (bimodală, trimodală etc., după numărul de valori mod). Distribuţiile unimodale conduc mai uşor la interpretări clare. Apariţia mai multor valori mod impune cercetătorului o ţinută ştiinţifică rezervată, fără afirmaţii categorice. Din astfel de situaţii se poate ieşi, eventual, printr-o transformare a scalei.

În reprezentările grafice, modul iese în evidenţă prin dreptunghiul cel mai înalt al histogramei sau prin sectorul de cerc cel mai extins al diagramei circulare.

Operaţiile efectuate asupra unei scale pot să schimbe valoarea mod. Utilizând scala IV a lui Bales pentru datele din tabelul 2.6, se obţine o nouă distribuţie, prezentată în tabelul 2.7. Se observă că valoarea mod este (ii) cu 95 de cazuri şi, mai mult, valoarea mod anterioară (A) nu a participat prin grupare la crearea noii categorii mod (ii). Deci prin grupare nu se respectă nici o regulă prestabilită de obţinere a noilor valori mod. Aceeaşi situaţie se constată şi la rafinarea unei scale. Dependenţa valorii mod de scala adoptată duce la relativizarea interpretării sale.

Tabelul 2.7. *Distribuţia intervenţiilor (scala IV a lui Bales)*

Categoria de comunicare	Număr
i	73
ii	95
iii	62

Analizând exemplele date mai sus, se constată că valoarea mod nu furnizează nici o informaţie despre repartizarea celorlalte date ale distribuţiei, fapt ce rezultă

de altfel şi din definiţia valorii mod. Forţând puţin lucrurile, să comparăm distribuţiile prezentate în tabelele 2.8 şi 2.9. Observăm că ambele au aceeaşi valoare mod (2) dar restul distribuţiei este esenţial diferit: în a doua distribuţie categoriile 1 şi 4 sunt insignifiante în comparaţie cu categoriile 3 şi 5 etc. Compararea celor două distribuţii prin intermediul doar al valorilor mod va duce la interpretarea falsă a unei similitudini între cele două distribuţii. Apare astfel necesitatea unui indicator care să reflecte şi caracteristicile împrăştierii datelor. Putem preciza acum că rezumatele statistice se divid în două categorii: **statistici privind tendinţa centrală** şi **statistici privind împrăştierea datelor.** Valoarea mod este o statistică de tendinţă centrală, singura suportată de scalele nominale.

Tabelul 2.8. *Eşantion A*

Categorie	Număr
1	8
2	20
3	8
4	8
5	8

Tabelul 2.9. *Eşantion B*

Categorie	Număr
1	2
2	20
3	14
4	2
5	14

În privinţa rezumării gradului de împrăştiere a datelor unei scale nominale prezentăm[10] **indicele de variaţie calitativă**. Alţi autori utilizează entropia, dar calculul este mai greoi, iar interpretarea este mai puţin specifică (prezentarea entropiei se va face ulterior).

Indicele de variaţie calitativă, notat în continuare IVC, compară variaţia *observată*, calculată pe baza datelor distribuţiei, cu variaţia *aşteptată*, calculată pentru distribuţia cu variaţie extremă. Formula este:

$$\text{IVC} = \frac{\text{variaţia observată}}{\text{variaţia aşteptată}} \times 100\%, \qquad \text{(II.1)}$$

unde atât variaţia observată, cât şi variaţia aşteptată sunt definite ca fiind *suma tuturor produselor posibile de câte două frecvenţe, fiecare pereche de frecvenţe fiind luată o singură dată*, iar distribuţiile implicate fiind cele precizate mai sus.

Astfel, dacă notăm efectivele claselor nominale cu n_1, n_2, ..., n_k, rezultă că variaţia calculată pe baza acestei distribuţii şi utilizată în formula 2.1 este:

$$\sum_{i \neq j} n_i \times n_j = n_1(n_2 + n_3 + \ldots + n_k) + n_2(n_3 + \ldots + n_k) + \ldots + n_{k-1}n_k.$$

Aplicând această formulă distribuţiei din tabelul 2.8, obţinem:

$$\text{variaţia observată} = (8\times20+8\times8+8\times8+8\times8) + (20\times8+20\times8+20\times8) + \\ + (8\times8+8\times8)+(8\times8) = 1024.$$

10. *Apud* Norman R. Kurtz, *op.cit.*, p. 69.

Distribuţia cu împrăştierea maximă este cea în care cele 52 de cazuri sunt repartizate cât mai uniform între categoriile scalei, deci distribuţia 11, 11, 10, 10, 10. Variaţia aşteptată este atunci:

$$\text{variaţia aşteptată} = (11\times11+11\times10+11\times10+11\times10)+(11\times10+11\times10+ +11\times10)+(10\times10+10\times10)+(10\times10)=1081.$$

Aplicând formula 2.1, obţinem pentru distribuţia din tabelul 2.8:

$$IVC=\frac{1024}{1081}\times100=9473\%,$$

ceea ce permite să afirmăm că împrăştierea datelor este apropiată de variaţia maximă, procentajul de 94,73 fiind foarte apropiat de 100%. Deci distribuţia corespunzătoare tabelului 2.8 este aproape uniformă, repartizarea datelor fiind suficient de echilibrată.

Procedând în mod analog, în cazul distribuţiei din tabelul 2.9 obţinem IVC = 88,07%, ceea ce arată, de asemenea, o distribuţie echilibrată. Comparând cei doi indici de variaţie calitativă, se poate afirma că distribuţia din tabelul 2.8 are un grad mai mare de concentrare a datelor, ceea ce corespunde şi intuiţiei.

Indicele de variaţie calitativă este utilizat totuşi doar pentru scale nominale cu relativ puţine categorii, situaţie de altfel uzuală. Este, de asemenea, de remarcat faptul că prin intermediul său se raportează o situaţie de fapt, observată, la o situaţie ideală, aşteptată, procedeu cu o anumită extindere şi importanţă în statistică, după cum se va vedea ulterior.

2.2. Scala ordinală

Scalele ordinale sunt cunoscute şi sub numele de **scale de ordine, scale de rang, scale ierarhice.** O scală ordinală permite ordonarea observaţiilor, persoanelor, situaţiilor de la mic la mare, de la simplu la complex, de la puţin la mult etc., atunci când în empiricul relativ se introduc, pe lângă relaţiile de echivalenţă, şi relaţii de ordine.

În cazul scalelor ordinale se poate stabili ierarhia a *n* gradaţii ale variabilei, dar nu se precizează (sau nu se poate preciza) valoarea diferenţei dintre două gradaţii.

2.2.1. Alcătuirea unei scale ordinale

Crearea unei scale ordinale presupune, cu prioritate, stabilirea unei relaţii de ordine între valorile posibile ale variabilei studiate.

Exemplificăm prin ierarhia nevoilor umane în concepţia lui Abraham Maslow.

Scala stabilită de acesta cuprinde, în ordine:

1. Nevoi fiziologice;
2. Nevoi de securitate;
3. Nevoi sociale, de apartenenţă la un grup;

4. Nevoia de stimă, de a fi apreciat de alţii;
5. Nevoia de a se realiza pe sine, de a concretiza prin acte şi performanţe potenţialul de talente şi cunoştinţe pe care o persoană crede că le posedă.

Ordinea prezentată este de la simplu la complex. Nu justificăm fundamentarea ştiinţifică şi străduinţele teoretice ale autorului de a funcţionaliza această scală. Dar, din momentul în care acceptăm să operăm cu ea, utilizăm caracteristicile unei scale ordinale.

Prin această scală putem să caracterizăm manifestările unei persoane în funcţie de gradul de saturare motivaţională. Dacă o persoană A este încadrată la nivelul 2, o persoană B la nivelul 3, o altă persoană C la nivelul 4, atunci putem face afirmaţia că persoana C are satisfăcute un număr mai mare şi mai complex de nevoi decât persoana B, care este în aceeaşi situaţie faţă de persoana A. Dar, caracteristic unei scale ordinale, nu putem preciza cu exactitate cu cât este mai complex C faţă de B sau B faţă de A şi nici nu putem compara cele două diferenţe (dintre C şi B, dintre A şi B) între ele.

Un alt exemplu este cel oferit de Irving Taylor (1959) privind clasificarea creativităţii după criteriul efectivităţii comunicării. Scala are următoarele 5 gradaţii[11], ordonate de la „îngust" la „extins":

1. Creativitate expresivă;
2. Creativitate productivă;
3. Creativitate inventivă;
4. Creativitate novatoare;
5. Creativitate emergentă.

Jean Piaget[12] oferă modelul unei scale ordinale după criteriul slab structurat — înalt structurat, prin care se evaluează stadiul inteligenţei la copil:

Tabelul 2.10. *Scala Piaget*

Stadiul	Denumire
1	exerciţiul reflexelor
2	primele adaptări dobândite şi reacţia circulară primară
3	reacţiile secundare şi procedeele destinate să prelungească spectacolele interesante
4	coordonarea schemelor secundare şi aplicarea lor la situaţii noi
5	reacţia circulară terţiară şi descoperirea de mijloace noi prin experimentare activă
6	inventarea de mijloace noi prin combinaţie mentală

11. Erika Landau, *Psychologie der Kreativität*, Ernst Reinhardt Verlag, München, Basel, 1969, p. 76.
12. J. Piaget, *Naşterea inteligenţei la copil*, Editura Didactică şi Pedagogică, Bucureşti, 1973.

În toate cele trei exemple de scale, operația de evaluare (măsurare) a unei persoane este deosebit de complexă și laborioasă, presupunând metode psihologice rafinate și practicieni experimentați, dar metodologia de clasare a unei persoane într-o asemenea scală face obiectul altor specialități din psihologie.

Putem sintetiza spunând că operația de realizare a unei scale ordinale presupune ca observațiile (categoriile) să poată fi diferențiate unele de altele, la fel ca în cazul scalelor nominale. Dar apare în plus ordonarea lor, ceea ce va conferi noi proprietăți măsurătorilor efectuate cu o scală ordinală.

2.2.2. Prezentarea datelor ordinale

Dacă se neglijează ordinea dintre categorii, datele obținute în urma utilizării unei scale de măsură ordinale sunt identice, din punct de vedere formal, cu datele nominale: denotă categoria în care au fost clasificate diferitele observații efectuate. Acestea vor fi raportate sub formă de frecvențe; prezentarea lor sub formă brută nu poate comunica aproape nimic, mai ales în cazul unui număr restrâns de categorii. Pe tot parcursul prelucrării este necesar, pentru simplificarea operațiunilor, să se efectueze o codificare a categoriilor scalei.

Simbolurile care se atribuie categoriilor scalei ordinale nu vor mai fi doar diferite, simple etichete, ci ordonări sau simboluri ordonate. Se preferă numere cu ordonare general acceptată. Astfel, se pot folosi codurile 1, 2, 3, care pot fi înlocuite prin 5, 9, 12 (păstrând ordinea naturală), dar în nici un caz nu pot fi substituite prin 4, 3, 7 (șir neordonat).

Precizăm încă o dată că aceste coduri (gradații ale scalei), care trebuie să realizeze doar ordonarea, nu poartă informații despre distanțele dintre gradații și nu indică precis cantitatea prezentă din respectiva caracteristică.

Toate noțiunile folosite pentru a explica funcționarea scalei nominale rămân valabile și în cazul scalei ordinale. Apar însă și noțiuni noi, condiționate de complexitatea mai mare a scalelor ordinale.

Datorită „înghețării" ordinii diviziunilor, are sens să vorbim de **forma distribuției**. Aceasta poate fi descrisă luând ca repere următorii termeni: simetrie – asimetrie, grad de asimetrie, asimetrie stânga – asimetrie dreapta, grad de aplatizare.

Cunoașterea formei unei distribuții este importantă prin aceea că oferă indicii asupra repartizării cazurilor în diverse categorii, permite corecții în cazul în care se utilizează modele teoretice și, în consecință, poate conduce la revizuiri ale scalelor utilizate etc.

Figurile 2.5, 2.6 și 2.7 ilustrează câteva situații specifice.

Stabilirea direcției asimetriei se face în funcție de partea existentă „în plus". În figura 2.6 se observă că eliminarea, ipotetică, a unei forme aproximativ simetrice din partea stângă lasă un „surplus" înspre dreapta; de aici denumirea de *asimetrie spre dreapta* sau *pozitivă*. Există și o justificare teoretică, dar este evident că aceasta are ca motivație tot observația anterioară.

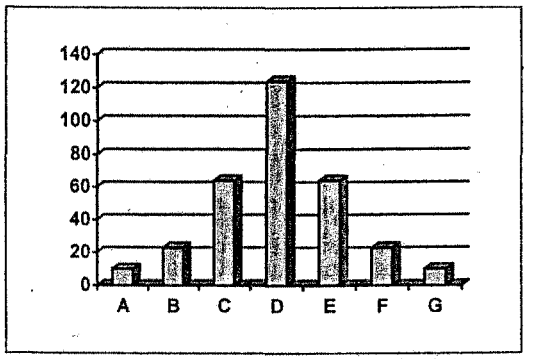


Figura 2.5. *Distribuție simetrică*

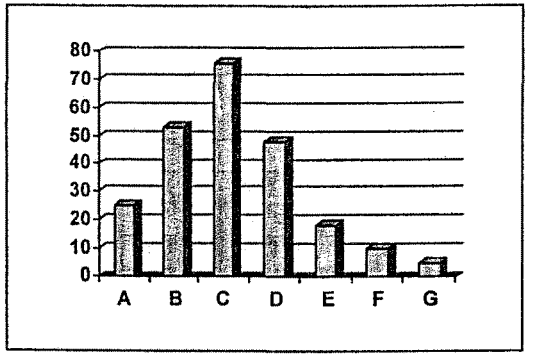


Figura 2.6. *Distribuție ușor asimetrică spre dreapta*

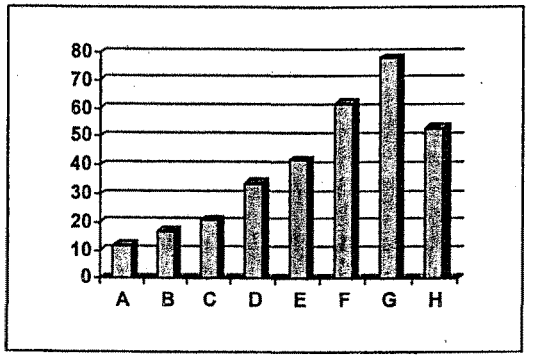


Figura 2.7. *Distribuție puternic asimetrică spre stânga*

Cunoașterea gradului de asimetrie poate conduce la concluzii de genul: eșantionul studiat nu este reprezentativ (lipsesc categorii sau observații dacă știm

că trebuie să obţinem o distribuţie simetrică, iar cea obţinută nu este astfel); realitatea conţine categorii în plus (dacă se utilizează un model teoretic simetric pentru analiza unei variabile despre care ştim că de obicei este asimetrică) etc.

Situaţia ilustrată în figura 2.8 este una de excepţie (majoritatea distribuţiilor acceptate sunt unimodale) şi este interpretată, de obicei, ca fiind „suprapunerea" a două distribuţii distincte. În asemenea cazuri se încearcă verificarea formării eşantionului studiat, a criteriilor de selecţie, a variabilelor urmărite etc. pentru a separa cele două populaţii statistice presupuse.

Precizăm că distribuţia exemplificată în figura 2.8 este unimodală în sensul strict al cuvântului (deoarece există o singură frecvenţă maximă), dar, în sens relativ, sunt considerate multimodale şi distribuţiile în care se pot distinge mai multe frecvenţe maxime locale (B, G din figură).

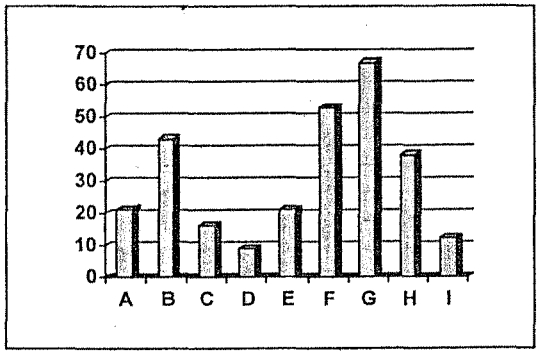


Figura 2.8. *Distribuţie bimodală asimetrică.*

Prezentarea sub formă de tabel a datelor ordinale conţine în plus, faţă de prezentarea datelor nominale, frecvenţele cumulate.

Frecvenţa cumulată pentru o categorie este definită ca fiind suma tuturor frecvenţelor *anterioare* plus frecvenţa curentă. Cuvântul „anterior" utilizat în definiţie se referă la sensul de parcurgere a categoriilor scalei ordinale. Dacă parcurgerea se face de la categoriile mici spre cele mari, se formează frecvenţe cumulate ascendent. Dacă parcurgerea se face de la categoriile mari spre cele mici, se formează frecvenţe cumulate descendent. Se pot prezenta sub formă de frecvenţe cumulate atât frecvenţele absolute, cât şi cele relative, astfel încât se poate vorbi despre frecvenţe absolute cumulate ascendent sau despre frecvenţe relative cumulate descendent etc.

Cumularea frecvenţelor îşi află justificarea în existenţa ordinii dintre categoriile scalei ordinale şi permite găsirea simplă a răspunsurilor la întrebări de genul: Câţi subiecţi au note mai mici sau egale cu 7? Câţi copii au nivelul de dezvoltare intelectuală cuprins în stadiile 1, 2, 3 şi 4 (după scala elaborată de Piaget)?

Exemplul 2.3. Să presupunem că, analizând produsele creative dintr-o anumită zonă geografică, pe o anumită perioadă de timp, s-a obţinut următoarea distribuţie de valori pe scala Taylor:

Tabelul 2.11. *Date privind produsele creative*

Nivel de creativitate	Frecvenţa absolută
1	75
2	52
3	20
4	6
5	4
Total	157

Calcularea frecvenţelor relative şi a celor cumulate este ilustrată în tabelul 2.12. Interpretarea numerelor marcate în tabel se face astfel:

- 147, din coloana frecvenţelor cumulate ascendent, este obţinut prin adunarea frecvenţelor absolute corespunzătoare categoriilor 1, 2 şi 3: 75 + 52 + 20. Acest număr ne arată numărul de produse care se încadrează în una dintre primele trei categorii.
- 82, din coloana frecvenţelor cumulate descendent, este obţinut prin cumularea frecvenţelor corespunzătoare categoriilor 2, 3, 4 şi 5. Rezultatul reprezintă numărul de produse care se încadrează în ultimele patru categorii.
- 93,6 constituie suma frecvenţelor relative ale primelor trei categorii şi conţine informaţia că 93,6% dintre produse se încadrează în primele trei categorii.

Tabelul 2.12. *Tabelul statistic complet pentru datele din tabelul 2.11*

Nivel	Frecvenţe absolute	Frecvenţe relative (%)	Frecvenţe absolute cumulate ascendent	Frecvenţe absolute cumulate descendent	Frecvenţe relative cumulate ascendent (%)	Frecvenţe relative cumulate descendent (%)
1	75	47,8	75	157	47,8	100,0
2	52	33,1	127	82	80,9	52,2
3	20	12,7	147	30	93,6	19,1
4	6	3,8	153	10	97,4	6,4
5	4	2,6	157	4	100,0	2,6
Total	157	100				

Reprezentările grafice ale datelor ordinale trebuie să reflecte proprietatea de ordonare specifică scalei, să denote faptul că scala de măsură este discretă şi să sugereze cât mai puţin luarea în considerare a distanţei dintre categoriile scalei. Din aceste considerente, sunt utilizate frecvent diagramele sub formă de bare verticale (cu aceeaşi formă ca şi în cazul datelor nominale). Nu sunt potrivite diagramele cu bare orizontale (pentru că ordinea este estompată). Atunci când se doreşte vizualizarea structurii setului de date, se poate utiliza totuşi forma de diagramă de structură.

Vizualizările pot să aibă drept obiect atât frecvențele absolute, cât și frecvențele relative. Dacă datele reprezentate sunt frecvențe cumulate, atunci acest fapt trebuie să fie indicat în mod explicit pentru a nu produce confuzii cu reprezentările de date necumulate.

Pentru exemplul anterior obținem astfel figurile 2.9 și 2.10. În prima dintre aceste figuri sunt utilizate frecvențele brute, iar diagrama sugerează ordinea categoriilor prin scrierea uzuală de la stânga la dreapta; dreptunghiurile sunt disjuncte pentru a arăta caracterul discret al datelor. Faptul că toate dreptunghiurile au aceeași lățime constituie aici doar un atribut estetic și nu se dorește sugerarea unui fapt care nu este adevărat, și anume acela că toate categoriile au o aceeași întindere (nu există noțiunea de distanță în categoriile ordinale). Trebuie totuși menționat faptul că o asemenea idee poate fi transmisă chiar fără a se dori acest lucru.

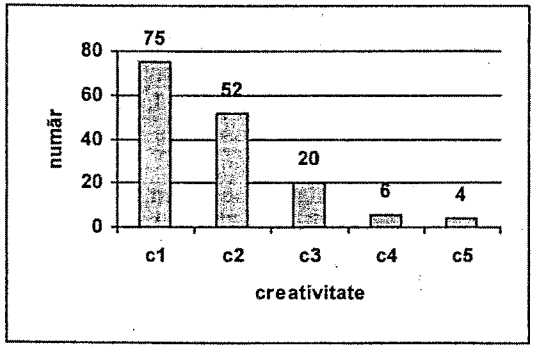


Figura 2.9. *Distribuția produselor creative (date absolute)*

În figura 2.10 sunt reprezentate frecvențele relative cumulate ascendent.

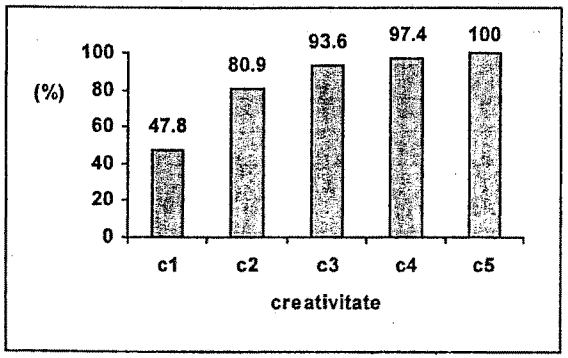


Figura 2.10. *Reprezentarea frecvențelor cumulate*

Se poate observa că, pentru a recepta informația prezentă într-un asemenea grafic, este necesar un anumit antrenament. Un astfel de grafic poate fi confundat

ușor, uneori în mod involuntar, cu o histogramă obișnuită. Din acest motiv se va specifica dacă datele reprezentate sunt cumulate sau nu. Informația conținută într-un asemenea grafic este deosebită prin aceea că se pot citi direct procentajele proprii unor grupuri de categorii (calculul algebric implicat este suficient de simplu pentru a necesita explicații).

2.2.3. Operații asupra scalelor ordinale

Operațiile introduse pentru scalele nominale își păstrează valabilitatea, cu precizarea că noile scale obținute prin condensare sau rafinare trebuie să fie gradate după regulile unei scale ordinale. De exemplu, dacă am cumula primele două categorii din scala Taylor, noua categorie nu poate avea ca etichetă un număr ≥ 3 (codul 3 și următoarele sunt deja atribuite categoriilor superioare). În asemenea cazuri se preferă renumerotarea (recodificarea) întregii scale.

O scală ordinală poate suporta orice transformare monotonă (care păstrează ordinea). Putem aduna astfel peste tot același număr, putem înmulți (împărți) cu același număr etc. Tabelul 2.13 conține exemplificări în acest sens.

Tabelul 2.13. *Transformări monotone în cadrul scalelor ordinale*

Gradația inițială	A	B	C	D
2	0	3	1	6
4	2	5	2	12
6	4	7	3	18
8	6	9	4	24
10	8	11	5	30

Coloana A se obține scăzând 2 din gradația inițială; coloana B se obține adunând 1 peste tot; coloana C se obține împărțind cu 2 toate valorile; coloana D se obține înmulțind cu 3 coloana inițială.

Se observă că, prin intermediul acestor transformări monotone (evident că nu sunt enumerate toate transformările posibile), numerele obținute pot alcătui gradațiile unei scale ordinale. Transformări de tip B, D sau similare oferă gradații mai „rarefiate" (care pot fi completate în eventualitatea unei rafinări a scalei). Prezentarea are însă și scopul de a accentua relativitatea unei codificări, singura cerință fiind aceea de a păstra ordinea categoriilor.

2.2.4. Statistici la nivelul unei scale ordinale

Rezumarea statistică a unei distribuții de date ordinale utilizează toate statisticile definite pentru scala nominală, cu completări de interpretare impuse de natura scalei ordinale. În afara acestora, se definesc statistici noi, unele dintre ele fiind doar prefigurarea simplificată a unor indicatori adecvați pentru scala de intervale.

Valoarea **mod** este definită drept categoria cu frecvența cea mai mare. Prin raportarea valorii mod se specifică gradația unde are loc o „aglomerare" a

observaţiilor. Acest indicator oferă cea mai bună predicţie în privinţa categoriei unei observaţii viitoare. Este evident că această predicţie este cu atât mai bună, cu cât frecvenţa valorii mod este mai mare decât celelalte frecvenţe ale distribuţiei.

Pentru exemplul 2.3 se obţine ca valoare mod categoria 1 (creativitate expresivă) având o frecvenţă egală cu 75, iar distribuţia fiind unimodală. Se poate aduce în discuţie poziţia valorii mod între gradaţiile scalei. În exemplul avut în vedere se observă că valoarea mod o constituie prima categorie, ceea ce exclude, oarecum, posibilitatea ca distribuţia să fie simetrică.

În comparaţiile care se realizează între diferite distribuţii se analizează atât valoarea, cât şi poziţia relativă a categoriei modale.

Dacă scala utilizată cuprinde suficient de multe categorii, iar numărul de observaţii este relativ restrâns, este posibil ca raportarea valorii mod să nu transmită informaţia corectă de tendinţă centrală a datelor. Să luăm în considerare exemplul unei distribuţii bazate pe o scală cu 30 de categorii (frecvenţele pentru categoriile 5...27 sunt toate egale cu 15):

categoria	1	2	3	4	...	28	29	30
frecvenţa	15	15	16	15	...	15	15	15

Aplicarea strictă a definiţiei conduce la afirmaţia că valoarea mod o constituie categoria 3. Este însă evident că distribuţia aproape uniformă a observaţiilor nu necesită raportarea unei valori mod. În asemenea cazuri, condensarea scalei poate produce rezultate situate mai aproape de spiritul definiţiei valorii mod.

Pentru aprecierea gradului de împrăştiere a datelor, se poate utiliza **indicele de variaţie calitativă** (IVC), ca alternativă la entropie, în cazul în care scala conţine relativ puţine categorii. Se recomandă ca aprecierea gradului de împrăştiere să se facă împreună cu precizarea formei distribuţiei (asimetrie, aplatizare).

Mediana, notată cu Me, este definită drept acea categorie a scalei de măsură care împarte observaţiile în două jumătăţi: jumătate dintre observaţii sunt clasate sub categoria mediană. Altfel spus, frecvenţa valorilor mai mici decât mediana este egală cu frecvenţa valorilor mai mari decât mediana.

Determinarea medianei în cazul unei distribuţii de date ordinale se poate dovedi un proces cu rezultat oarecum imprecis. Unii autori[13] introduc, pentru date ordinale, chiar noţiunea de categorie cvasimediană, cu scopul de a surprinde aproximările necesare în acest caz. Abia în cazul datelor de interval mediana poate fi determinată în mod exact.

Cazul datelor negrupate

Notăm cu $X_1, X_2, \ldots, X_N$ datele observate şi cu $X_{(1)}, X_{(2)}, \ldots, X_{(N)}$ şirul obţinut prin ordonarea crescătoare (sau descrescătoare) a datelor observate. În cazul în care N este impar, atunci:

13. Henry Rouanet; Brigitte Le Roux, *Statistique en sciences humaines. Exercices & Solutions*, Dunod, Paris, 1995, p. 4.

$$Me = X_{\left(\frac{N+1}{2}\right)},$$

cu alte cuvinte, valoarea din mijlocul şirului ordonat este mediana distribuţiei observate. Dacă N este par, atunci nu mai există o valoare centrală şi se poate alege drept mediană:

$$Me = X_{\left(\frac{N}{2}\right)} \quad \text{sau} \quad Me = X_{\left(\frac{N+2}{2}\right)}.$$

Observaţie. Unii autori consideră că mediana este, în cazul unui număr par de date ordinale, media aritmetică a valorilor centrale. Pentru datele ordinale, operaţiunea de mediere nu are sens (există locul I, locul II, dar nu există locul I şi jumătate). Introducerea unui rang intermediar are sens doar dacă se consideră că scala ordinală poate fi transformată într-o scală continuă, ceea ce ridică semne de întrebare asupra construcţiei riguroase a scalei de măsură. Dacă variabila cercetată poate fi măsurată pe mai multe tipuri de scale, se va alege de la început scala cea mai complexă, care oferă cele mai complete prelucrări.

Din păcate, de cele mai multe ori în aplicaţii concrete, valorile centrale, indiferent de paritatea lui N, nu sunt unice. În asemenea situaţii, fixarea valorii mediane se va face încercând să se respecte cât mai bine spiritul definiţiei, iar rezultatul poate să se depărteze suficient de mult de ceea ce simţul comun înţelege prin punctul de mijloc.

Să luăm în considerare şirul de valori ordinale:

12, 5, 8, 19, 11, 9, 7

care prin ordonare devine

5, 7, 8, 9, 11, 12, 19.

Se observă că şirul ordonat are 7 termeni, deci mediana ar trebui să fie a patra valoare ordonată (valoarea din mijloc). Cum această valoare este unică, nu există nici o urmă de îndoială şi obţinem Me = 9.

Dacă şirul de valori este:

12, 5, 9, 19, 9, 7, 9, 9, 10, 12,

care, ordonat, devine

5, 7, 9, 9, 9, 9, 10, 12, 12, 19,

atunci N = 10, fiind deci par, iar mediana trebuie fixată între valorile a cincea şi a şasea. Cum ambele valori date sunt egale cu 9, se pare că, fără îndoială, mediana este 9. Interpretând însă această valoare, ar trebui să afirmăm că jumătate dintre

valorile observate se află sub valoarea 9, ceea ce este fals, pentru că doar două valori sunt mai mici decât 9. În această situație, încercăm să analizăm cea mai bună alegere având în vedere și următoarea modalitate, 10. Se observă că sub 10 sunt șase dintre observații (60%), pe când sub 9 se află doar două dintre observații (20%). Este evident că alegerea valorii Me = 10 este mai aproape de sensul definiției medianei.

Se poate afirma astfel că, în cazul datelor ordinale, ordonarea valorilor observate oferă doar un indiciu asupra localizării medianei. Determinarea finală poate rezulta în urma unei negocieri între valorile candidate.

Cazul datelor grupate

În această situație, se pleacă de la o distribuție de frecvențe (cazul anterior fiind un caz particular, cu cvasitotalitatea frecvențelor egale cu 1). Determinarea medianei este facilitată de calculul prealabil al frecvențelor cumulate, absolute sau relative.

În cazul tabelului statistic 2.12 (care conține date grupate pe variante de variație), cum N = 157, trebuie determinată categoria scalei care depășește jumătate dintre observații, adică N/2 = 78. Din coloana frecvențelor absolute cumulate ascendente se obține că a șaptezeci și noua observație este în categoria 2 (care are singură 52 de observații). Teoretic, ar trebui ca din cele 52 de observații proprii să reținem doar 4, ceea ce, evident, nu conduce la nimic semnificativ, deoarece ne situăm în aceeași categorie (de remarcat încă o dată caracterul necontinuu al datelor ordinale). Din încercarea de balansare, explicată la datele negrupate, se observă că nivelul 2 depășește 75 de observații, iar nivelul 3 depășește 127 de observații. Mai aproape de jumătatea numărului de observații este nivelul 2, deci Me = 2.

Același calcul poate fi efectuat utilizând frecvențele relative cumulate ascendente. Rezultă că nivelul 2 depășește 47,8% dintre observații (deci aproape 50%), iar nivelul 3 depășește 80,9% dintre observații (mult mai depărtat de 50%). Se va alege Me = 2.

În cazul în care se procedează la gruparea categoriilor prin operațiunea de condensare a scalei, calculul medianei se va efectua pe scala transformată, chiar dacă raportarea se face în funcție de vechile categorii.

Quantile. Definiția medianei, asociată esențial cu proporția de 1/2 (simbolizând jumătatea), poate fi generalizată prin luarea în considerare a altor proporții. Se obține astfel o clasă de indicatori numite quantile (sau **cvantile, cuantile**). O quantilă reprezintă o categorie a scalei de măsură care depășește o proporție precizată de observații.

Notăm atunci cu Q_p quantila de ordin p, unde $p \in (0, 1)$ este o proporție fixată, definitorie pentru quantila definită.

Interpretarea acestui indicator este:

$$\frac{\text{numărul de observații mai mici decât } Q_p}{\text{numărul total de observații}} = p.$$

Se observă atunci că mediana este quantila de ordin 1/2: $\text{Me} = Q_{\frac{1}{2}}$.

Pentru datele ordinale, determinarea quantilelor poate să se abată uneori de la strictețea definiției, în sensul precizat la determinarea medianei.

Pentru a determina quantila de ordin p, se parcurg următoarele etape:

- se ordonează observațiile;
- se determină rangul quantilei dorite prin $k = [pN] + 1$, unde p este ordinul quantilei, N este numărul total de observații și $[pN]$ notează partea întreagă a lui pN (cel mai mare întreg care nu depășește pe pN).

De exemplu, pentru N = 125 de observații, quantila de ordin 2/10 este valoarea observației de rang k = [(2/10) × 125] + 1 = 26.

Pentru facilitarea interpretării, se exprimă uneori proporțiile sub formă procentuală, deci se poate vorbi de quantila de (ordin) 15% sau quantila de (ordin) 75% etc.

Prin faptul că definesc jaloane sub care se găsesc proporții ale numărului de observații, quantilele oferă o imagine a gradului de împrăștiere a observațiilor și sunt utilizate adesea pentru a fixa poziția unei observații în raport cu mediana distribuției.

Această imagine este întregită dacă se utilizează sisteme de quantile, precizându-se astfel împrăștierea pe întregul domeniu al datelor observate. Tabelul următor conține cele mai frecvente sisteme de quantile utilizate în psihologia aplicată:

Nr.crt.	Număr de părți egale	Număr quantile	Denumire
1	3	2	Trecilă
2	4	3	Quartilă
3	5	4	Quintilă
4	6	5	Sextilă
5	7	6	Septilă
6	8	7	Octilă
7	9	8	Nonilă
8	10	9	Decilă
9	100	99	Centilă

Quartile

Sistemul determinat de quantilele de ordine 1/4, 2/4 = 1/2, 3/4 realizează (teoretic) o împărțire a observațiilor în patru părți egale. Aceste quantile se numesc **quartile** și sunt notate, cel mai adesea, prin Q_1, Q_2 și, respectiv, Q_3. Prin urmare:

$$Q_1 = Q_{\frac{1}{4}},\ Q_2 = Q_{\frac{1}{2}},\ Q_3 = Q_{\frac{3}{4}}.$$

În mod echivalent, se pot utiliza termenii: quartila întâi (sau quartila de 25%), a doua (sau quartila de 50%) şi, respectiv, a treia (sau quartila de 75%). Este evident că mediana este identică cu quartila a doua.

Se observă că doar pentru o repartiţie uniformă (toate categoriile au aceeaşi frecvenţă) quartilele sunt egal depărtate (ca număr de categorii) între ele.

Prin poziţia pe care o au quartilele distribuţiei, se poate aprecia împrăştierea datelor şi se pot efectua comparaţii între eşantioane sau între variabile chiar de esenţă diferită.

Astfel, pentru un elev care are 7,55 la matematică şi 8,20 la engleză este greu să se precizeze disciplina unde elevul are aptitudini sporite. Aprecierea globală că 8,20 este mai mare decât 7,55 nu este corectă, deoarece se compară valori diferite (sistemele de notare nu sunt atât de standardizate încât să putem spune că doi profesori dau note exact la fel, cunoştinţele sunt de naturi diferite etc.). O rezolvare a problemei (nu singura) poate fi încadrarea notelor respective în sistemul de quartile determinate pentru clasa, şcoala etc. elevului. Dacă 7,55 este mai mare decât Q_3 (din sistemul pentru matematică), iar 8,20 se situează între quartilele a doua şi a treia (în sistemul pentru engleză), putem spune că elevul este mai dotat pentru matematică, întrucât depăşeşte la matematică mai mult de 75% dintre elevi, în timp ce la engleză depăşeşte doar mai mult de 50% (dar mai puţin de 75%) dintre elevi.

Intervalele dintre quartile (termenul de interval este oarecum impropriu pentru o scală ordinală) sunt denumite **intervale interquartile** şi, uneori, prin abuz de limbaj, sunt denumite tot quartile. Situaţia este înţeleasă din context sau precizată; confuzia nu produce, de altfel, erori majore de interpretare.

Decile

Quantilele de ordine 1/10, 2/10, 3/10, 4/10, 5/10, 6/10, 7/10, 8/10, 9/10 poartă denumirea de **decile**, deoarece realizează o împărţire a eşantionului în zece părţi de volume egale. Ele sunt definite drept decila întâi (decila de 10%), decila a doua (decila de 20%), ..., decila a noua (decila de 90%) şi se pot nota Q_1, Q_2, Q_3, ..., Q_9, dacă se precizează că este vorba despre decile:

$$Q_1 = Q_{\frac{1}{10}},\ Q_2 = Q_{\frac{2}{10}},\ldots,Q_9 = Q_{\frac{9}{10}}.$$

Datorită cotelor procentuale mai mici, rezultă că prin sistemul de decile se stabileşte o structură mai fină de apreciere a împrăştierii datelor.

Intervalele dintre decile sunt denumite **intervale interdecile** sau tot decile.

Centile (percentile)

Quantilele de ordine 1/100, 2/100, 3/100, ..., 98/100, 99/100 poartă denumirea de **centile** (sau **percentile**), deoarece realizează o împărţire a eşantionului într-o sută de părţi de volume egale. Aceste valori sunt definite drept centila întâi (centila

de 1%), centila a doua (centila de 2%) etc. Interpretarea centilelor este imediată, ele fiind asociate cu procentele: centila de 35% depăşeşte 35% din populaţie etc.

Pentru o distribuţie de date grupate, se poate utiliza următoarea formulă (cu observaţia că aplicarea ei la date ordinale nu este justificată decât într-o oarecare măsură), formulă care defineşte **rangul centilic** al categoriei i a scalei de măsură ordinale:

$$c_i = \frac{100}{N}\left(Fc\uparrow - \frac{f_i}{2}\right),$$

unde c_i este rangul centilic, $Fc\uparrow$ numărul de observaţii mai mici decât scorul de interes (frecvenţa absolută cumulată ascendent), iar f_i reprezintă frecvenţa absolută a categoriei i.

Rangul centilic oferă o estimare a procentajului de observaţii depăşite de categoria implicată.

> **Observaţie.** Utilizarea decilelor şi, mai ales, a centilelor este justificată atunci când volumul eşantionului este suficient de mare. Nu are sens să calculăm, de exemplu, centilele unei distribuţii de 50 de elemente.

2.3. Scala de intervale

Scala de intervale are o complexitate mai mare, în comparaţie cu scala ordinală, oferită de proprietatea că are sens luarea în considerare a distanţelor dintre categoriile scalei. Prin aceasta, operaţia de măsurare a observaţiile pe o scală de intervale determină cantitatea exactă de caracteristică existentă în observaţii. Astfel, scala de intervale oferă, pe lângă ordinea categoriilor (la fel cu scala ordinală), distanţa dintre categorii (inexistentă la scala ordinală). Această proprietate nouă face ca datele experimentale obţinute pe o asemenea scală (numite date de interval) să suporte prelucrări statistice mai complexe.

Utilizatorul unei astfel de scale are nevoie să compare intervalele determinate prin respectiva scală, adică diferenţa dintre numerele pe care scala le atribuie claselor de observaţii.

2.3.1. Alcătuirea unei scale de intervale

Realizarea unei scale de intervale presupune parcurgerea mai multor etape.

În primul rând, este necesar să se fixeze o operaţie experimentală care face posibilă definirea a ceea ce se înţelege când cineva spune că distanţa sau diferenţa dintre două aspecte (caracteristici) cercetate este egală cu distanţa dintre alte două aspecte (caracteristici).

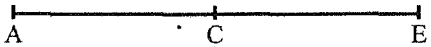
O dată realizată prima etapă, se vor putea atribui numere caracteristicilor studiate, în aşa fel încât la două distanţe experimentale egale să corespundă două diferenţe numerice egale. Această cerinţă privind codificarea este importantă în toate prelucrările numerice ale datelor de interval.

Exemplul cel mai la îndemână al unei scale de intervale este acela al unei rigle gradate cu care se pot măsura lungimi. Se observă faptul că există o ordine bine precizată între gradaţiile riglei şi se poate afirma imediat că, de exemplu, diferenţa sau distanţa dintre 10 cm şi 15 cm este aceeaşi cu diferenţa dintre 28 cm şi 33 cm.

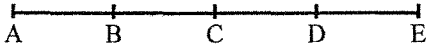
Deşi realizarea egalităţii între două caracteristici este relativ simplă în alte domenii unde se efectuează măsurători (lungimi, temperaturi, mase etc.), acest deziderat poate fi deosebit de dificil în ştiinţele sociale. În psihologie se reflectă situaţia respectivă prin aceea că au trebuit să fie utilizate adesea nu numai constatări experimentale, ci şi anumite postulate.

Singurele operaţii care permit, în psihologie, definirea directă a două intervale sunt aşa-numitele operaţii de echipartiţie. Un subiect supus la o asemenea operaţie este numit „estimator" sau „judecător". Unor astfel de subiecţi li se poate cere să claseze frazele descriind grade diferite de reuşită şcolară sau pe cele care induc diferite grade de anxietate, să aprecieze opinii inegal favorabile în diferite privinţe, în aşa fel încât intervalele dintre ele să pară egale. Experienţele arată că reuşitele estimatorilor în astfel de încercări lasă de dorit în privinţa preciziei.

De exemplu, cercetările care urmăresc găsirea unui punct de mijloc între două senzaţii auditive sau vizuale nu au dus la rezultate concludente. Cercetarea lui Delbœuf cere unui pictor să realizeze un gri care să fie la fel de îndepărtat de un eşantion de alb şi de un eşantion de negru (deci un punct de mijloc). Nu trebuie să cădem în eroarea de a crede că operaţia este pur tehnică şi constă în amestecul în proporţii egale de alb şi negru. Fondul problemei este aici de a realiza o echidistanţă subiectivă, senzorială, faţă de efectele produse de imaginea de alb şi cea de negru. Experienţele lui Gage (1934) privesc aceeaşi problemă a echipartiţiei intervalelor în domeniul senzaţiilor auditive. Un subiect trebuia să subdividă un interval senzorial AE în două intervale egale printr-un stimul C.



Operaţia era repetată pentru intervalele AC şi CE, obţinându-se punctele B, D, care înjumătăţesc intervalele AC şi, respectiv, CE:



Dacă operaţia de înjumătăţire a unui interval senzorial ar fi posibilă pentru subiect, atunci mijlocul intervalului BD ar trebui să coincidă cu C. Experienţele lui Gage au demonstrat însă depărtări semnificative de la această cerinţă.

H. Piéron explică imprecizia unui asemenea demers prin ambiguitatea sarcinii propuse subiectului, căruia i se cere să ia în considerare intensitatea senzaţiilor, pe când el are în vedere, în mod obişnuit, intensitatea stimulilor fizici care îi suscită senzaţiile.

Prezentarea experienţelor anterioare nu duce totuşi la concluzia că în psihologie nu este posibilă alcătuirea unei scale de intervale. În realizarea pragurilor diferenţiale, limitele clasei de diferenţiere se stabilesc printr-un postulat: *în momentul în care o limită este percepută ca fiind mai mică sau mai mare în 50% dintre cazuri*. Dificultăţile realizate în operaţia de ordonare în intervale egale între anumite observaţii pot fi evitate dacă realizăm o nouă procedură, constând în a declara că toate proprietăţile astfel stabilite au valabilitate doar în cadrul unei convenţii de limbaj. Un exemplu foarte cunoscut este postulatul lui Fechner, care afirmă că toate intervalele determinate prin scala ordinală a senzaţiilor sunt egale. Maurice Reuchlin[14] susţine că este lipsită de sens dezbaterea dacă această convenţie este „adevărată" sau „falsă". Orice rezultat obţinut cu ajutorul unei astfel de scale nu are sens decât în cadrul acestei convenţii. Astfel, legea lui Fechner după care senzaţia creşte proporţional cu logaritmul excitaţiei are valoare doar dacă se acceptă postulatul acestuia.

O dată ce s-a stabilit egalitatea a două intervale (prin convenţie sau nu), devine posibilă operarea unor divizări repetate ale fiecărui interval în subintervale (asemănător divizării metrului în decimetri, a decimetrilor în centimetri etc.). O scală de interval este, prin urmare, divizibilă la infinit. Din această cauză, o scală de intervale este numită **scală continuă**, spre deosebire de scalele nominale şi ordinale care sunt **scale discrete**.

Faptul că diferitele persoane au praguri senzoriale absolute şi diferenţiale dincolo de care nu mai pot sesiza diferenţele nu anulează rezultatele statistice obţinute pe baza proprietăţilor teoretice ale unei scale de intervale.

Scalele utilizate în biometrie sunt, evident, scale de intervale. Astfel, caracteristici precum vârsta, greutatea, înălţimea, timpul de reacţie sunt măsurate pe asemenea scale.

Deşi acceptarea unor postulate privind egalitatea unor intervale pare a rezolva, din punct de vedere teoretic, problema alcătuirii unei scale de intervale, realizarea practică a acestora în domeniul cuantificării unor trăsături psihice este deseori imperfectă. Însăşi problema stabilirii unor convenţii presupune un înalt „simţ" al adecvării la domeniul investigat. Ilustrările prezentate în continuare nu fac excepţie.

1. Durata medie a erorilor la testul de asociere şi disociere manuală DM1 elaborat de Lahy

Unul dintre criteriile de apreciere a randamentului la proba DM1 îl constituie durata medie a erorilor. Deoarece datele au fost obţinute în unităţi de timp (zecimi de secundă), ele sunt măsurate pe o scală de intervale. Culegerea datelor brute într-un test DM1 îngreunează foarte mult aprecierea, deoarece ele se află în contextul altor date provenind din teste cu amplitudini foarte diferenţiate. Din acest motiv se înlocuieşte notarea valorii observate cu notarea intervalului în care se înscrie valoarea observată.

14. M. Reuchlin, *op.cit.*

Tabelul 2.14. *Intervalele de grupare pentru testul DM1*

Intervalul (clasa)	Limita inferioară	Limita superioară
I	19,60	22,2
II	17,00	19,5
III	14,40	16,9
IV	11,70	14,3
V	9,10	11,6
VI	6,50	9,0
VII	3,90	6,4
VIII	3,18	3,8
IX	mai mic decât 3,18	

Înregistrarea intervalelor (în locul valorilor brute) nu elimină caracterul continuu al variabilei.

Observaţie. Nu trebuie să se confunde semnificaţia matematică a noţiunilor inferior/superior cu semnificaţia psihologică de randament superior sau inferior. Un subiect care obţine o valoare de 21,5 se apropie, matematic, de limita superioară, dar, în aceeaşi clasă, el este cotat mai slab, din punct de vedere psihologic, ca performanţă de test. Această confuzie poate să apară în multe prelucrări de date psihologice.

Pe lângă funcţionalitatea de clasare a unui subiect într-un interval, prin care se evaluează performanţa subiectului în raport cu un grup, culegerea datelor sub forma anterioară permite, în cazul unei cercetări de grup, obţinerea imediată a unui tabel de frecvenţe.

2. Chestionarul de personalitate 16PF (R.B. Cattel)

Un anumit tip de răspuns poate primi 2, 1 sau 0 puncte. Valoarea numerică a unui factor este rezultatul însumării punctajelor la anumiţi itemi. Doi subiecţi care au aceeaşi notă la un anumit factor pot avea repartiţiile punctajelor elementare foarte diferite între ele. Faptul că sunt apreciaţi identic din punct de vedere numeric constituie rezultatul unei convenţii.

Se pot elabora diferite *pattern*-uri de răspuns, care pot face obiectul de interes al unor analize de rafinament.

2.3.2. Prezentarea datelor de interval

Colecţiile de date obţinute în cadrul unei scale de intervale sunt mult mai nuanţate decât cele de la scalele nominale sau ordinale. Acest fapt se datorează caracterului continuu al scalei. Deoarece se măsoară cantitatea exactă de caracteristică prezentă (în limitele discutate anterior), observaţiile pot fi foarte diferite între ele, şi nu doar în limitele unui număr finit de categorii de clasificare.

Observaţie. Datorită limitelor umane sau ale aparatelor utilizate în măsurare, poate fi creată impresia că sunt utilizate un număr finit de gradaţii ale scalei. De exemplu, măsurând înălţimi între 1,50 m şi 2,00 m, se pot obţine doar 51 de numere distincte dacă precizia este la centimetru. Acest fapt nu înseamnă că doi subiecţi care sunt înregistraţi ca având 1,75 m au exact aceeaşi înălţime: unul dintre ei ar putea avea un milimetru în plus, fapt care nu este avut în vedere. Deci relativa finitudine a numărului de gradaţii utilizate nu provine din natura scalei (care este continuă), ci din gradul de precizie acceptat în cercetarea respectivă.

Exemplul 2.4. Măsurând timpul de reacţie, în miimi de secundă, pentru 200 de pacienţi, se obţin, în ordinea determinărilor, valorile de mai jos. Practic, se pot obţine 200 de valori diferite între ele. Sub această formă de adunătură haotică, datele oferă un conţinut informaţional aproape nul.

201 190 222 318 163 192 310 163 247 254 331 241 258 298 160
249 240 263 207 227 242 181 240 185 257 232 339 334 281 241
249 257 204 183 171 298 283 253 240 275 213 189 237 338 281
241 183 225 233 335 329 254 257 261 276 228 199 245 324 287
298 256 272 266 197 241 251 276 229 315 284 241 278 203 230
316 312 242 257 197 298 284 247 260 279 213 312 291 288 243
177 293 309 273 260 218 197 305 300 243 253 279 207 168 301
297 244 259 216 233 297 304 250 178 304 284 248 272 263 218
228 285 286 254 258 288 179 228 238 288 243 251 184 212 239
235 266 254 275 223 204 244 255 258 225 236 189 243 263 223
245 187 200 248 256 219 239 233 259 200 219 223 231 250 257
249 251 294 202 231 224 248 254 271 265 206 254 255 210 259
262 271 236 228 260 218 264 268 269 244 222 223 224 337 326
326 238 294 229 257

Înainte de prelucrarea statistică propriu-zisă, datele trebuie să fie supuse unor operaţiuni de ordonare şi concentrare. Aceleaşi date, dacă sunt prezentate după ce sunt ordonate crescător, oferă informaţiile într-o formă mult mai accesibilă:

160 163 163 168 171 177 178 179 181 183 183 184 185 187 189
189 190 192 197 197 197 199 200 200 201 202 203 204 204 206
207 207 210 212 213 213 216 218 218 218 219 219 222 222 223
223 223 223 224 224 225 225 227 228 228 228 228 229 229 230
231 231 232 233 233 233 235 236 236 237 238 238 239 239 240
240 240 241 241 241 241 241 242 242 243 243 243 243 244 244
244 245 245 247 247 248 248 248 249 249 249 250 250 251 251
251 253 253 254 254 254 254 254 254 255 255 256 256 257 257
257 257 257 257 258 258 258 259 259 259 260 260 260 261 262
263 263 263 264 265 266 266 268 269 271 271 272 272 273 275
275 276 276 278 279 279 281 281 283 284 284 284 285 286 287
288 288 288 291 293 294 294 297 297 298 298 298 298 300 301
304 304 305 309 310 312 312 315 316 318 324 326 326 329 331
334 335 337 338 339

Se pot citi direct timpii extremi: timpul minim = 160, timpul maxim = 339, ordinul de mărime al timpilor înregistrați, diferența dintre timpii extremi, numărul variantelor de variație etc. Ca argument în favoarea unei astfel de operații se poate aduce, şi nu ca unul oarecare, confortul psihologic al cercetătorului care analizează datele. Se poate simți însă şi sub această formă o accesibilitate redusă la înțelegerea globală a situației. Este foarte greu de receptat situația datelor multiple (cum ar fi 241 care apare de 5 ori etc.), unde sunt „aglomerate" valorile etc. Acest aspect este cu atât mai accentuat, cu cât creşte numărul de valori (în definitiv, 200 de observații nu reprezintă o cercetare vastă). Este necesară, prin urmare, concentrarea datelor. Această operație se poate desfăşura în două etape (vezi şi capitolul privind ordonarea datelor):

- marcarea valorilor multiple şi raportarea lor sub formă de frecvențe;
- gruparea datelor în intervale de grupare.

Prima modalitate constă în construirea tabelului următor (tabelul este doar schițat; într-o prelucrare concretă se completează toate cele aproximativ 110 linii necesare, dacă se marchează doar modalitățile prezente):

Timp înregistrat	Frecvența
160	1
163	2
168	1
...	...
254	6
...	...
338	1
339	1

Modalitățile prezente în acest tabel sunt încă foarte multe, ceea ce nu produce simplificarea necesară.

A doua posibilitate de concentrare a informației este oferită de definirea unor intervale de grupare şi raportarea valorilor sub formă de frecvențe de interval. Se obține astfel un tabel de forma următoare:

Interval	Frecvența
160-179	8
180-199	14
200-219	20
220-239	32
240-259	56
260-279	26
280-299	22
300-319	12
320-339	10

Sub această formă de prezentare, informația poate fi receptată mai deplin, cresc posibilitățile de operare a datelor, dar trebuie avut în vedere faptul că s-a pierdut o parte din informația inițială (de exemplu, din ultimul tabel nu rezultă exact care sunt cele 8 valori din primul interval: mai aproape de 160, de 170, de 179 etc.). Această pierdere de informație se justifică totuşi prin avantajul oferit de simplificare, ca şi prin aceea că o astfel de uniformizare a datelor (în cadrul fiecărui interval) poate să reducă unele erori posibile în cadrul operației tehnice de măsurare.

Observație. Prelucrarea electronică a datelor a redus necesitatea grupării datelor pentru prelucrare, impunându-se însă necesitatea grupării pentru raportarea rezultatelor.

2.3.2.1. Limite de grupare

Prin natura scalei de intervale, trebuie să facem distincție între două categorii de limite de grupare. Analizând gruparea prezentată în tabelul anterior, este evident că limitele intervalelor au avantajul că sunt uşor de înțeles (numere întregi, din 20 în 20 etc.) şi de asimilat. Asemenea limite sunt numite **limite de raportare** şi sunt alese astfel încât să fie acceptate sociologic (prin tradiție, prin convenție etc.) de către psihologi.

Dezavantajul unor asemenea limite este acela că intervalele definite nu acoperă în întregime domeniul datelor continue. Astfel, o valoare de 179,2 nu aparține nici unui interval. Din acest motiv, trebuie să se definească alte limite, care să permită clasarea fiecărei observații din domeniul luat în considerare. Pentru stabilirea noilor limite se porneşte de la regula de rotunjire a numerelor. Numim **limite exacte** ale unui interval acele valori pentru care valorile intermediare aparțin, prin rotunjire, intervalului avut în vedere.

Astfel, intervalul cu limitele 160 şi 179 (limitele de raportare) are limitele exacte 159,5 şi 179,5. Matematic, intervalul este exprimat [159,5;179,5). Este evident că orice valoare între aceste limite aparține, după rotunjire, intervalului inițial.

În practică, este posibilă şi ordinea inversă de stabilire a limitelor: mai întâi se stabilesc limitele exacte şi apoi cele de raportare, care permit redactarea clară a rapoartelor de cercetare.

Orice prelucrare a datelor grupate, care implică limitele de grupare, este efectuată prin luarea în considerare a limitelor exacte, deşi rezultatele pot fi raportate cu ajutorul limitelor de raportare.

Notăm, pentru nevoile prezentării, cu l_i limita inferioară exactă şi cu l_s limita superioară exactă a unui interval. Lungimea intervalului este obținută în acest caz prin scăderea limitelor: $l = l_s - l_i$.

2.3.2.2. Centrele intervalelor

Prin gruparea datelor se pierde informația primară, astfel încât nu se mai cunoaşte valoarea măsurată a fiecărei observații. În calcule sunt însă necesare valorile

individuale ale observațiilor. Din această cauză, fiecare valoare, prezentă într-o distribuție de date grupate doar prin intermediul frecvențelor, trebuie să fie aproximată cât mai exact.

Numim **centrul unui interval**, notat cu *c* (urmat, eventual, de indicele intervalului), valoarea situată la mijlocul intervalului respectiv. Ca formulă de calcul se poate utiliza, evident:

$$c = l_i + l/2 \text{ sau } c = (l_i + l_s)/2,$$

unde notațiile sunt cele introduse în 2.3.2.1 pentru limitele exacte și lungimea intervalului. Astfel, centrul primului interval din exemplul anterior este:

$$c = (159{,}5 + 179{,}5)/2 = 169{,}5.$$

Se acceptă aproximarea fiecărei valori individuale prin centrul intervalului (clasei) căruia îi aparține. Astfel, fiecare dintre cele 8 observații, care sunt prezentate ca aparținând primului interval de grupare, se va considera egală cu centrul intervalului, mai exact cu 169,5. Cu alte cuvinte, dacă nu cunoaștem valorile inițiale, putem considera că am efectuat 8 observații cu rezultatul 169,5.

Eroarea maximă posibilă datorată acestui tip de aproximare este egală cu jumătatea lungimii intervalului (*l/2*). Aproximarea prin centrul intervalului satisface criteriul celor mai mici erori, în sensul că utilizarea oricărui alt tip de aproximare poate să conducă la erori mai mari, fapt ușor de verificat. Evaluarea erorii maxime susține și ideea că un interval de grupare cu lungime mare produce erori mai mari.

2.3.2.3. Reprezentări grafice ale datelor de interval

În mod uzual, datele de interval sunt reprezentate grafic după operațiunea de grupare a lor. Sub formă grupată, tabelul corespunzător este similar unui tabel de distribuție de frecvențe ordinale și, prin urmare, reprezentarea grafică poate fi similară reprezentării datelor ordinale: diagrame cu bare verticale pentru frecvențe brute (cumulate sau nu) sau frecvențe relative (cumulate sau nu). Este însă recomandat ca, pentru a evidenția continuitatea datelor de interval, barele verticale să fie unite între ele. Se obține astfel o **histogramă** a datelor. De exemplu, tabelul de frecvențe din exemplul anterior produce reprezentarea următoare:

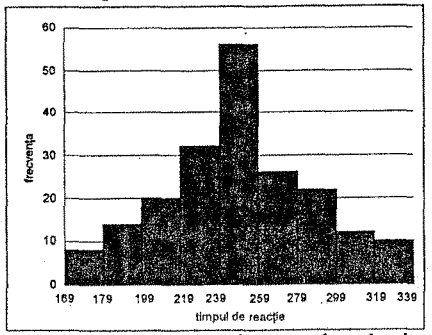


Figura 2.11. *Histograma frecvențelor absolute*

O histogramă, pe lângă o imagine directă a repartiției frecvențelor, oferă informații privind forma distribuției și simetria repartiției. Este dificil de apreciat gradul de aplatizare, deoarece, manevrând unitățile de măsură pe axa Oy a diagramei, se pot obține, pentru același set de date, configurații mai mult sau mai puțin aplatizate.

Datorită complexității crescute a datelor de interval, sunt posibile și reprezentări grafice specifice, care nu sunt adecvate datelor discrete. Prin unirea punctelor de mijloc ale laturilor superioare dintr-o histogramă, se obține **poligonul frecvențelor**. După cum se prelucrează frecvențe absolute sau relative, cumulate sau nu, se poate vorbi de poligonul frecvențelor brute, poligonul frecvențelor relative cumulate etc. Două exemple sunt prezentate în continuare:

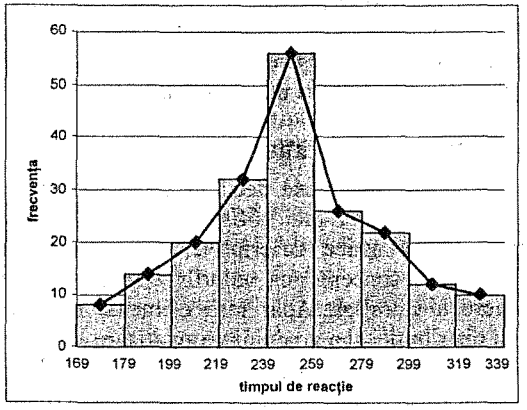


Figura 2.12. *Poligonul frecvențelor brute*

Prin figurarea poligonului frecvențelor brute, apare mai clară forma distribuției (mai ales dacă nu sunt desenate și coloanele histogramei).

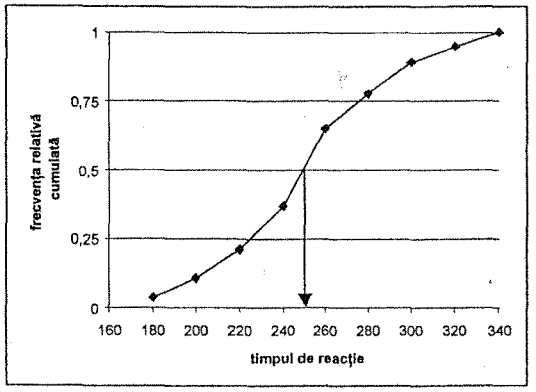


Figura 2.13. *Poligonul frecvențelor relative cumulate*

În reprezentarea frecvențelor relative cumulate s-a respectat indicația unor autori[15] de a nu utiliza punctele situate la mijloacele intervalelor, ci punctele având ca abscise limitele superioare exacte ale intervalelor. Aceasta presupune apoi unirea prin segmente a punctelor din dreapta sus ale laturilor dreptunghiurilor din histograma frecvențelor relative cumulate. Interpretarea acestui tip de poligon este:

- ordonata corespunzătoare unei valori x arată proporția de observații mai mici sau egale cu x;
- abscisa corespunzătoare unei ordonate p este valoarea care depășește o proporție p de observații. Pe grafic se poate citi astfel că la o proporție de 0,5 (adică 50%) corespunde valoarea aproximativă 250, cu semnificația că 50% dintre observații au avut înregistrat un timp mai mic sau egal cu 250 (250 este astfel o estimație a medianei repartiției);
- diferența proporțiilor p_1 și p_2, corespunzătoare la două valori x_1 și x_2, arată proporția observațiilor cu valori măsurate între x_1 și x_2.

2.3.3. Statistici pentru date de interval

Scala de intervale fiind mai complexă decât scalele discrete, datele măsurate pe asemenea scale pot suporta cele mai sofisticate tehnici statistice. După logica generală a prelucrărilor statistice descriptive, și pentru datele continue sunt urmărite aspectele privind concentrarea valorilor (tendința centrală a datelor) și gradul de împrăștiere a valorilor. Apar însă elemente teoretice noi, cum ar fi aprecierea simetriei distribuției și a gradului de aplatizare.

Datorită volumului important de informații, în acest capitol se vor prezenta doar indicatorii de tendință centrală, urmând ca indicatorii de împrăștiere, simetrie și aplatizare să fie studiați într-un capitol distinct.

2.3.3.1. Indicatorii de tendință centrală

Calculul indicatorilor rezumativi pentru o distribuție de date continue este influențat de modul de prezentare a datelor. Datele negrupate pot să producă rezultate nesemnificative, în timp ce datele grupate pot să producă rezultate aproximative (datorită erorilor care apar din procesul de grupare). În prezentarea fiecărui indicator statistic, se vor preciza limitele de aplicare și interpretare în funcție de gruparea datelor.

Valoarea mod

Definiția este aceeași cu cea introdusă la datele nominale sau ordinale. În cazul datelor negrupate, valoarea mod se definește ca fiind valoarea cea mai frecventă. Aplicarea strictă a definiției în acest caz poate să producă un rezultat fără utilitate practică, datorită numărului important de valori posibile. Astfel, în șirul de valori

12, 15, 15, 21, 23, 25, 32, 45, 52, 57

15. N.R. Kurtz, *op.cit.*

se observă că valoarea mod este 15 (are frecvența maximă = 2), dar nu se poate spune că în șirul dat are loc o aglomerare a valorilor în jurul acestui rezultat.

În cazul datelor grupate, valoarea mod se definește ca fiind centrul intervalului care are frecvența maximă. Exemplul 2.4, luând în considerare tabelul datelor grupate, conduce astfel la o valoarea mod egală cu 249,5 (mijlocul intervalului 240-259, care are frecvența maximă). Dacă datele sunt reprezentate sub formă de histogramă, valoarea mod este centrul intervalului reprezentat prin dreptunghiul de înălțime maximă. În poligonul frecvențelor absolute, valoarea mod este abscisa punctului maxim.

Mediana

Introdusă ca punctul de mijloc al distribuției de valori, mediana poate fi determinată, în cazul datelor negrupate, după o metodă similară celei prezentate pentru datele ordinale:

- se ordonează valorile $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$;
- dacă n este impar, atunci mediana este valoarea din mijloc, adică $x_{\left(\frac{n+1}{2}\right)}$;
- dacă n este impar, atunci mediana este media aritmetică a celor două valori din mijlocul seriei ordonate, adică

$$Me = \frac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n+2}{2}\right)}}{2}.$$

În cazul datelor grupate, se va determina mai întâi intervalul median, adică intervalul care conține mediana. Este de remarcat faptul că din însăși procedura de grupare are loc etapa de ordonare a valorilor (intervalele de grupare sunt ordonate prin natura lor). Prin urmare, intervalul median este intervalul care conține observația din mijlocul seriei ordonate și poate fi identificat din șirul frecvențelor cumulate drept intervalul (clasa) a cărui frecvență cumulată depășește jumătatea numărului de observații. Mediana este atunci calculată după formula

$$Me = l_i + l \times \frac{\frac{n}{2} - f_{i-1}\downarrow}{f_i}.$$

unde: l_i este limita inferioară exactă a intervalului median, l este lungimea intervalului median, n este numărul total de observații, $f_{i-1}\downarrow$ este frecvența cumulată pentru intervalul anterior intervalului median, f_i este frecvența (necumulată) a intervalului median.

Formula este, de fapt, una de interpolare liniară prin care se adaugă la limita inferioară exactă a intervalului median o lungime proporțională cu numărul de observații care mai sunt necesare pentru atingerea observației din mijlocul seriei ordonate.

Unii autori[16] recomandă o formulă simplificată: mediana este egală cu centrul intervalului median. Este evident că diferența dintre cele două rezultate este acceptabilă în cazul în care lungimea intervalului median este mică. În plus, formula de interpolare dată presupune că observațiile sunt distribuite uniform în intervalul median, ceea ce nu se verifică întotdeauna.

Pentru exemplul 2.4 formăm mai întâi tabelul cu frecvențele cumulate ascendent.

Interval	f_i	$f_i\downarrow$
160-179	8	8
180-199	14	22
200-219	20	42
220-239	32	74
240-259	56	130
260-279	26	156
280-299	22	178
300-319	12	190
320-339	10	200

Cum n = 200, observațiile de la mijloc sunt observațiile cu rangurile 100 și 101. Se observă că acestea aparțin intervalului 240-259 pentru care frecvența cumulată depășește prima dată 100. Avem, prin urmare, l_i =239,5; l = 20; n = 200; $f_{i-1}\downarrow$ = 74; f_i = 56. Aplicând formula de calcul al medianei, obținem 248,79. Putem interpreta astfel că jumătate dintre observații sunt situate sub această valoare.

Dacă am utiliza ca mediană mijlocul intervalului median, am obține valoarea 249,5, care nu este foarte depărtată de valoarea obținută anterior.

Media aritmetică

Cea mai utilizată valoare rezumativă pentru caracterizarea tendinței centrale a datelor este media aritmetică, pe care o calculăm după formula uzuală:

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n},$$

unde n este numărul de observații, iar x_1, x_2, ..., x_n reprezintă seria de valori observate. Formula aceasta poate fi aplicată doar în cazul datelor negrupate pe intervale (doar atunci avem valorile individuale observate).

În cazul datelor grupate, fiecare valoare este aproximată prin centrul intervalului căruia îi aparține și, prin urmare, formula devine:

$$\bar{x} = \frac{f_1c_1 + f_2c_2 + \ldots + f_kc_k}{f_1 + f_2 + \ldots + f_k} = \frac{f_1c_1 + f_2c_2 + \ldots + f_kc_k}{n},$$

16. N. Rancu; L. Tövissi, *Statistică matematică cu aplicații în producție*, Editura Academiei, București, 1963.

unde k este numărul de intervale de grupare, f notează frecvențele, c notează centrele intervalelor, iar n este numărul de observații. Formula dă în fapt media ponderată a centrelor intervalelor, ponderile fiind frecvențele intervalelor.

Media aritmetică este punctul de echilibru al distribuției, în sensul că suma abaterilor de la medie pentru observațiile mai mici decât media este egală cu suma abaterilor de la medie pentru observațiile mai mari decât media. Acest rezultat se exprimă matematic prin:

$$(x_1 - \bar{x}) + (x_2 - \bar{x}) + (x_3 - \bar{x}) + \ldots + (x_n - \bar{x}) = 0.$$

Se mai spune că media aritmetică joacă rolul de centru de greutate al datelor.

În procesul de calcul, se poate recurge la transformări prealabile ale datelor observate, astfel încât calculele să se efectueze mai ușor. Transformările frecvente sunt:

$$y_i = x_i - A$$

și

$$y_i = A \times x_i,$$

prima transformare semnificând reducerea ordinului de mărime al datelor inițiale prin extragerea unei cantități constante A, iar a doua producând modificarea ordinului de mărime prin multiplicare cu o constantă A. De exemplu, pentru o serie de date între 100 și 200 se poate scădea 100 din toate valorile, calculele ulterioare efectuându-se pe valori mai mici. Pentru o serie de valori între 0,001 și 0,010 se pot înmulți toate valorile cu 1000. Media aritmetică a datelor inițiale se regăsește, corespunzător, prin formulele:

$$\bar{x} = A + \bar{y}$$

și

$$\bar{x} = \frac{\bar{y}}{A}.$$

În final se poate spune că media și mediana sunt cele mai utilizate statistici prin care se estimează tendința centrală a unei distribuții. Pentru distribuțiile perfect simetrice, valorile celor doi indicatori sunt identice. Pentru distribuțiile asimetrice, media tinde să se deplaseze către valorile extreme, spre dreapta sau spre stânga, în sensul asimetriei, în timp ce mediana rămâne mai apropiată de ramura „scurtă" a distribuției. În asemenea cazuri, mediana poate oferi o imagine mai bună a centrului. Situația se poate remedia, pentru medie, dacă se exclud valorile aberante.

Folosirea mediei în cercetarea psihologică este frecventă în trei tipuri de situații, dintre care unele pot fi rezolvate și prin luarea în considerare a medianei sau a valorii mod:

- situarea unei observații într-o distribuție. Exemplu: precizarea faptului dacă o notă este superioară sau inferioară mediei;

- compararea unui grup de observaţii cu un alt grup. Exemplu: în cadrul proceselor de învăţare, care comportă o serie de încercări succesive, pot fi lăsate intervale de timp (învăţare distribuită) sau nu (învăţare masată). Se constată, în general, că rezultatul unui grup de subiecţi în învăţarea distribuită este superior rezultatului în învăţarea masată. Dacă cineva consideră că această diferenţă este prea mare pentru a putea fi explicată doar prin variaţii fortuite, aceasta va conduce la o explicaţie psihologică;
- situaţia în care observaţiile au fost supuse la una sau mai multe surse de variaţie sistematică. Se poate calcula valoarea medie asociată la fiecare dintre stările sursei respective. Această medie permite descompunerea variaţiilor observate în mai multe abateri (care se numesc „efecte") şi, astfel, analizarea importanţei relative a diferitelor surse sistematice sau fortuite. O asemenea analiză este esenţială în explicarea fenomenelor observate.

Problema calculului mediilor apare şi în cazul analizei de itemi. Itemul este o componentă informaţională elementară a unui test sau chestionar psihologic, care, în majoritatea cazurilor, se prezintă dihotomic din punct de vedere cantitativ. Analiza de itemi constituie o operaţiune strict necesară în cazul în care dorim să ne asigurăm că testul îndeplineşte o serie de exigenţe funcţionale. Să luăm două exemple de modalităţi de prezentare a itemilor în probe psihologice:

	Test 1			
Nr.crt.	1	2	3	4
1	0	1	0	0
2	1	1	1	0
3	1	1	0	1
4	0	0	1	1
5	1	0	0	0
6	1	0	0	0
7	0	1	0	1
8	1	0	1	1
9	1	1	0	0
10	1	0	1	1
11	0	1	0	0
f_i	7	6	4	5
p_i	0,63	0,54	0,36	0,45
q_i	0,37	0,46	0,64	0,55

Test 2			
1	2	3	4
2	5	4	4
3	1	4	2
1	1	2	1
1	3	2	4
2	5	4	1
3	2	4	4
1	3	1	2
2	2	3	3
3	4	5	4
3	2	4	4
2	1	4	3
23	29	37	36
0,41	0,52	0,67	0,65
0,59	0,48	0,33	0,35

În cazul primului exemplu, avem situaţia de examen a 11 subiecţi la un test cu 4 itemi (exemplul este teoretic, deoarece un test trebuie să aibă cel puţin 20 de itemi pentru a putea fi numit test, după părerea psihometricianului englez Paul Kline[17]). La fiecare item se poate primi valoarea 1 (soluţie corectă la problema pusă de item

17. P. Kline, „La construction des tests", în J.R. Beech; Leonora Harding, *Tests, mode d'emploi. Guide de psychométrie*, Les Éditions du Centre de Psychologie Appliquée, Paris, 1994, p. 103.

sau răspuns la un chestionar de personalitate, care scoate în evidenţă trăsătura psihică ce trebuie diagnosticată) sau 0 (soluţie incorectă sau răspuns care nu evidenţiază trăsătura psihică cercetată). La al doilea test, răspunsul are o distribuţie cantitativă polihotomică, scorurile itemilor variind de la 1 la 5 (situaţie întâlnită la unele subscale ale testului WISC şi, mai ales, la chestionarele de atitudini).

În primul caz, f_i indică punctajul total obţinut de toţi subiecţii la itemul i, p_i (numit indice de dificultate) desemnează proporţia de răspunsuri corecte şi, totodată, media răspunsurilor corecte la itemul i (situaţie întâlnită doar la notarea dihotomică a răspunsurilor). În acest caz, p_i se confundă cu $\bar{x}_i$ (adică media itemului i). Deci $p_i = f_i /11$ (11 fiind numărul total de subiecţi). În cazul al doilea, cu acelaşi număr de itemi şi acelaşi număr de subiecţi, p_i, adică proporţia de rezolvare a sarcinilor, se obţine din împărţirea sumei scorurilor, f_i (suma tuturor valorilor de la itemul 1 de la cel de-al doilea test), la 55, adică la suma totală care s-ar fi obţinut dacă toţi subiecţii ar fi avut maximum de punctaj. Se observă că, în acest caz, p_i nu se mai confundă cu $\bar{x}_i$, deoarece valoarea celui din urmă rezultă din formula $\bar{x}_i = f_i /11$. În cazul primului item de la al doilea test, putem concretiza $\bar{x}_i = 23/11 = 2{,}09$, în timp ce $p_i = 23/55 = 0{,}41$.

Exemplificarea prezentată este şi un avertisment pentru a nu se confunda orice tip de raportare cu o medie.

2.4. Scala de raport sau scala de proporţii

Această scală de măsură posedă ca note distincte existenţa unei origini naturale şi precizarea clară a semnificaţiei unităţii de măsură, ceea ce face posibilă compararea raporturilor dintre gradaţiile scalei. Scala de raport a fost denumită de S.S. Stevens *scală de interval cu origine raţională*. În acest caz, raporturile nu sunt alterate printr-o transformare de similaritate (o transformare liniară în care b = 0). Aceste scale sunt puţin utilizate în psihologie, deoarece majoritatea mărimilor nu-i satisfac axiomele de bază. Conform acestora, dacă a putut fi găsită o operaţie empirică prin care să se satisfacă în întregime axiomele lui Campbel, devine posibil să ştim dacă, între 4 măsuri ale acestei mărimi, A, B, C şi D, sunt verificate sau nu relaţiile următoare :

$$A - B = B - C = C - D \text{ şi } A/B = C/D.$$

Pe scala de raport, numerele oferă nu numai informaţii asupra ordinii de rang a obiectelor şi a mărimii relative a diferenţelor, ci şi despre relaţiile rapoartelor. Când numerele 2, 4, 7 şi 9 reprezintă gradul de reliefare a unei anumite caracteristici pentru obiectele A, B, C şi D într-o scală de raport, putem spune că B este dublul caracteristicii pe care o are A, C este de 3,5 ori mai mult faţă de caracteristica A, D de 2,5 ori mai mult decât B ş.a.m.d. Măsurarea presupune cunoaşterea (existenţa) punctului 0 (originea), precum şi a unui pas scalar identic

de-a lungul întregii scale. Se cunoaşte direcţia continuumului, există paşi egali ai scalei şi se poate raporta exact fiecare poziţie individuală a unui continuum la un punct de referinţă (originea scalei). Cele mai multe instrumente de măsură pe care le utilizăm în situaţii cotidiene oferă date pe o scală de raport, scală care este de fapt o extensie a scalei de interval. Cînd o anumită caracteristică a unui obiect este măsurată pe această scală, putem să precizăm gradul de reliefare şi să-l exprimăm printr-o valoare absolută. Valorile individuale pot fi precizate fără a cunoaşte valoarea altor obiecte pe aceeaşi scală. O scală de proporţii poate fi stabilită pentru un atribut care, la prima vedere, nu prezintă posibilitatea de absenţă totală, atunci când se poate demonstra existenţa unei relaţii funcţionale între valorile acestui atribut sau fenomen care satisface proprietatea de adunare (de unde existenţa lui 0 absolut).

Prima caracteristică a scalei de proporţii rezidă în posibilitatea de a cunoaşte distanţa dintre origine şi valoarea unui individ pe un continuum definit după un atribut dat. Amplitudinea absolută a unei cantităţi de atribut este un dat cunoscut şi este posibilă cunoaşterea poziţiei unui scor individual, fără să fie necesară situarea altor indivizi sau a altor obiecte pe continuum. Numerele reprezintă deci în mod real gradul de prezenţă a unui atribut la un individ sau la un obiect. Scala de proporţii este cea care permite informaţiile cele mai complete.

Dacă se multiplică sau se divizează scorurile unei scale de proporţii printr-o constantă, scala conservă proprietăţile sale, ordinea indivizilor şi a obiectelor rămâne aceeaşi, proporţiile între intervale şi între scoruri rămân neschimbate, iar punctul 0 este totdeauna prezent. Dacă se adună sau se scade o constantă la rezultate, natura scalei va fi modificată: ordinea obiectelor şi a indivizilor rămâne aceeaşi, dar proporţiile între scoruri vor fi modificate şi va dispărea 0 absolut. Practic, se revine la scala de interval. Se poate face o remarcă în acest sens: cu cât mai precise sunt informaţiile furnizate printr-o scală, cu atât mai mică este libertatea de a transforma scala fără a modifica informaţiile originare.

O scală ordinală poate să sufere orice transformare monotonă fără ca informaţiile relative primare, pe care ea le produce, să fie modificate. La o scală de interval, posibilităţile de transformare sunt mai restrânse. Transformările permise la scala de proporţii trebuie să se limiteze la multiplicări sau la diviziuni cu ajutorul unei constante, dacă dorim ca informaţia să fie identică. Postulatele subiacente modelului scalei de proporţii autorizează utilizarea tuturor operaţiilor algebrice (+, –, ×, :), atât în intervale, cât şi în scoruri. Posibilitatea de a utiliza cele 4 operaţii algebrice autorizează recursul la cele mai puternice metode ale analizei matematice şi ale statisticii, incluzând în plus toate operaţiile permise în scale mai puţin complexe, deja menţionate.

Astfel, când se corectează un test şi când se adună numărul de itemi corect rezolvaţi prin respondent, se obţine un număr situat pe o scală de proporţii. Dar, dacă se utilizează acest număr pentru a indica poziţia respondentului pe o scală care măsoară un atribut, numărul respectiv nu mai poate fi interpretat după regulile scalei de proporţii. Unui număr nul de răspunsuri corecte nu-i corespunde o absenţă

totală a atributului măsurat. Atributele psihologice nu permit să se aplice procedura de stabilire a unui relativ funcţional între două atribute în care unul prezintă posibilitatea de absenţă totală. Aşadar, scalele de proporţii nu pot să convină măsurării atributelor psihologice.

I.P. Vasilescu[18] susţine ideea potrivit căreia scalele Weber-Fechner sau Zapan, utilizate la măsurarea intensităţii senzaţiei în funcţie de intensitatea stimulului, satisfac exigenţele scalelor de raport.

În privinţa axiomelor se consideră că sunt respectate toate cele specifice numerelor reale. Aplicarea unei constante de deplasare nu este permisă, întrucât există un punct 0 natural.

Operaţiile statistice permise sunt cele care pornesc de la valorile absolute. Dintre mărimile statistice care pot fi determinate amintim media geometrică şi coeficientul de variaţie sau de variabilitate. Procedeele de verificare şi corelaţiile valabile sunt aceleaşi ca la nivelul anterior al scalelor de interval.

Există opinia că putem vorbi de utilizarea scalei de raport doar în privinţa măsurării timpului de reacţie în cercetarea caracteristicilor psihologice. Cercetarea timpului de reacţie este revendicată şi de alte discipline ştiinţifice, dar nu trebuie uitat faptul că în psihologie nu interesează atât timpul de reacţie, cât fenomenul psihic care se derulează într-un anumit timp. O parte din autorii care se ocupă de acest subiect consideră că valorile rezultate pot fi apreciate după cerinţele unei scale de interval.

Posibilitatea construirii scalelor de raport în psihologie preocupă cercetătorii de mai bine de un secol, realizările experimentale vizând mai ales domeniul senzorial. Primele tentative de a determina intensitatea fizică a stimulului suscitând la un subiect o senzaţie apreciată ca dublă în comparaţie cu alta aparţin lui J. Merkel şi au avut loc în anii 1888 şi 1889. La puţin timp, în anul 1892, G.S. Fullerton şi J.M. Cattel realizează experienţe în care o senzaţie trebuie să fie dublată sau, din contra, redusă la jumătate. Asemenea experienţe au fost abandonate până în anii '30, când au fost reluate din nevoi practice ale acusticienilor[19]. Interesul ştiinţific viza construirea unei scale care să permită să cunoaştem de câte ori un sunet de intensitate fizică de natură cunoscută apare mai intens subiectiv unui examinat, în comparaţie cu un alt sunet de intensitate fizică necunoscută. Domeniile psihofizicii şi teoriei măsurii în psihologie ar fi beneficiat astfel în acest sens. Tendinţa de a construi scale de raport în psihologie a cunoscut diferite forme. S.S. Stevens le-a clasat astfel:

- metode de apreciere a rapoartelor – subiectului îi sunt prezentaţi doi stimuli şi el se pronunţă imediat asupra valorii numerice a raportului între senzaţiile corespondente;

18. I.P. Vasilescu, *Statistică informatizată pentru ştiinţe despre om*, Editura Militară, Bucureşti, 1991, p. 19.
19. Vezi Maurice Reuchlin, *Les méthodes quantitatives en psychologie*, Presses Universitaires de France, Paris, 1962, pp. 40-50.

- metode de producere a rapoartelor – subiectul ajustează el însuşi un stimul în aşa fel încât raportul senzaţiilor pe care le suscită în comparaţie cu o senzaţie constantă să ia o valoare fixată în avans;
- metode de apreciere a mărimilor – subiectului i se cere să furnizeze o estimaţie numerică directă a fiecăreia dintre senzaţiile coresponzând unei serii de stimuli;
- metode de producere a mărimilor – subiectului i se cere să regleze un stimul variabil, în aşa fel încât senzaţia să ia o valoare care să-i pară a corespunde unui număr dat[20].

Dacă datele primite sunt considerate coerente, este uşor să construim scale ale senzaţiilor şi să determinăm eşaloanele corespondente unor intensităţi fizice ale stimulului. Se atribuie arbitrar valoarea 1 senzaţiei produse printr-un stimul, valoarea 2 senzaţiei produse de un stimul care pare dubla primei senzaţii ş.a.m.d. Adepţii acestui tip de scală sunt expuşi la două tipuri de obiecţii: unele de ordin logic şi altele de ordin experimental. Metoda a fost criticată chiar în şedinţele Comitetului de la „British Association for the Advancement of Science", instituţie ştiinţifică deosebit de prestigioasă. Argumentul logic central pare să fie următorul:

> Aserţiunea „senzaţia A este dublul senzaţiei B" este lipsită de sens, atâta timp cât adunarea empirică nu a putut fi definită. Subiectul nu va folosi şi nu va înţelege o astfel de expresie dacă nu este familiarizat cu mărimi fundamentale, precum lungimea şi greutatea, pentru care adunarea a fost definită. Dacă cineva spune că o lungime A este dublul unei lungimi B, e posibil să descrie o suită de operaţii fizice care să-i permită definirea a ceea ce el înţelege prin aceasta. Nu acelaşi lucru se poate petrece în domeniul sonic, în care nu se poate ataşa o semnificaţie metrică judecăţii subiectului supus la astfel de experienţe.

Din punct de vedere experimental, dacă se iau în considerare abaterile medii pentru un acelaşi subiect, în experienţele de acest tip se constată că ele depăşesc adesea 10 decibeli, anumite deviaţii individuale arătând deviaţii mai mari de 50 decibeli.

T.W. Reese a încercat, în 1943, să verifice scalele de raport pentru alte mărimi decât senzaţiile. El le-a utilizat pentru a măsura dificultatea subiectivă a fiecărei componente a unui test. O astfel de tentativă poate reuşi cu condiţia ca rezultatele experimentale să manifeste o coerenţă suficientă. El a cerut subiecţilor, care au învăţat două serii de cifre, să spună dacă dificultatea pe care au avut-o la memorizarea celei de-a doua serii, în comparaţie cu prima memorizare, le-a părut pe jumătate sporită sau redusă. Acelaşi procedeu este folosit şi pentru cuvintele unui test de vocabular. Dar coerenţa rezultatelor obţinute în operaţiile fundamentale de egalizare a raporturilor subiective nu este suficientă pentru a se putea afirma cu certitudine că această operaţie este posibilă empiric, cu un grad rezonabil de precizie. Îngustimea domeniului explorat prin aceste metode nu permite să se spună

20. Vezi Maurice Reuchlin, *op.cit.*, p. 41.

în mod sigur că ele constituie soluţia generală a problemelor puse de procesul de măsurare în psihologie.

În domeniul teoriei măsurării din ştiinţele umaniste se înfruntă acerb două şcoli, una vizând domeniul măsurării „slabe" (numită şi şcoala fundamentalistă şi având ca principali reprezentanţi pe S.S. Stevens, S. Siegel şi D.R. Saunders), iar cealaltă – domeniul măsurării „puternice" (care îi are ca principali protagonişti pe Q. McNemar, F.M. Lord, N. Anderson, W.I. Hays şi J.C. Nunnally). Controversele dintre cele două şcoli nu par să se fi atenuat până acum.

Fundamentaliştii susţin că trebuie respectat principiul izomorfismului între realitate şi sistemele matematice utilizate pentru a o reprezenta. În consecinţă, ei refuză să utilizeze, în măsurarea atributelor psihologice, tehnici statistice specifice scalelor de intervale egale. În centrul acestei polemici se situează postulatul egalităţii unităţilor. Celor care au experienţă în domeniul psihometric le este evident faptul că este dificil, chiar imposibil, să se considere că un atribut psihologic se poate manifesta în unităţi egale. Pentru fundamentalişti, demonstrarea acestei egalităţi constituie o condiţie esenţială pentru a putea utiliza tehnici statistice puternice. Partizanii şcolii statisticilor „puternice" susţin că acestea din urmă se aplică la numere, şi nu la atributele înseşi, deci caracteristicile reale ale scalelor de măsură nu au influenţă asupra tehnicilor statistice[21]. Ca atare, testul statistic răspunde la chestiunea pentru care este destinat să răspundă, indiferent dacă scala este ordinală sau de intervale egale. Argumentul a făcut obiectul demonstraţiei empirice că obţinerea unei scale de intervale egale nu constituie un preambul la utilizarea unui test parametric, în scopuri de inferenţă statistică. În urma acestei demonstraţii empirice, s-a stabilit că, în statistica inferenţială, tehnicile statistice puternice pot foarte bine să convină datelor ordinale, reprezentanţii şcolii statisticilor puternice acceptând că punctul de vedere restrictiv al fundamentaliştilor este pertinent când cineva se preocupă de determinarea unei statistici într-un scop descriptiv.

Opinia cea mai răspândită în rândul cercetătorilor avizaţi este aceea că nu există unităţi corecte sau veritabile pentru o scală de măsură a proporţiilor. Clasificarea unei scale va fi mai curând o problemă de convenţie şi de utilitate[22]. Convenţiile stabilite în ştiinţele umane permit în mod obişnuit tratarea datelor ca şi când ele ar satisface postulatul de egalitate a intervalelor. Tehnicile puternice de analiză statistică sunt, prin convenţie, compatibile cu datele privind atributele umane (incluzând şi atributele psihologice), cel puţin atunci când datele nu au fost obţinute iniţial printr-o procedură de „punere în rang", rezultând dintr-o scală de măsură ordinală.

Valorile acestui tip de scală sunt numerice şi pot fi supuse la operaţii aritmetice elementare.

Ca şi în fizică, trebuie stabilită o distincţie netă între scalele cu zero absolut şi zero convenţional. Scalele de temperatură, indiferent că sunt Fahrenheit sau

21. J.J. Bernier; B. Pietrulewicz, *La psychométrie*, Gaëtan Morin (éd.), Montréal, Paris, Casablanca, 1997.

22. Pentru consideraţii în această privinţă, *ibidem*, pp. 32-33.

Celsius, nu au un zero absolut. Fără zero absolut, este posibilă măsurarea distanţelor sau a diferenţelor, dar aceste valori individuale sau rapoartele lor nu sunt aditive. Scalele care comportă un zero absolut sunt scalele de proporţii. Cu ajutorul acestui tip de scală se elaborează segmentele în funcţie de lungimea lor, intervalele de timp în funcţie de durata lor şi sunetele în funcţie de intensitatea lor.

Capitolul 3

INDICATORI DE MĂSURĂ A ÎMPRĂŞTIERII

Valoarea informaţională a unui indicator statistic poate fi apreciată din trei puncte de vedere diferite:

- **istoric** – ce s-a întâmplat la un moment dat într-o situaţie dată;
- **comparativ** – cum se raportează situaţiile similare unele faţă de altele;
- **predictiv** – ce putem spune despre evoluţia viitoare a unui fenomen, bazându-ne pe cunoaşterea evoluţiei anterioare.

Să luăm notele obţinute de elevii A şi B, într-un an şcolar, la un acelaşi obiect de studiu:

elev A: 7, 8, 6, 9, 7, 8, 8, 7, 9, 8;
elev B: 5, 6, 9, 10, 5, 7, 9, 8, 10, 8.

Ambii elevi au media egală cu 7,70. Din acest motiv, se poate aprecia că ei au acelaşi nivel de pregătire. Acest fapt este însă valabil doar dacă se analizează situaţia din punctul de vedere istoric sau din cel comparativ (în anul 1997, elevii A şi B au avut aceeaşi medie, 7,70). Dacă suntem interesaţi de prognoza performanţelor şcolare ale celor doi elevi, situaţia este total diferită. Este evident că la o examinare ulterioară ne vom aştepta ca elevul A să fie apreciat cu o notă între 7 şi 8 (deci apropiată mediei lui), deoarece el este mult mai constant, în timp ce elevul B poate lua practic orice notă. Din punctul de vedere predictiv, valorile celor două medii nu mai coincid. Ceea ce produce diferenţa dintre seriile de date este doar gradul lor de împrăştiere: notele elevului A sunt mai „adunate", în timp ce notele elevului B sunt mai „împrăştiate". Analiza anterioară arată că, pe lângă indicatorii de tendinţă centrală, este necesar să se introducă indicatori ai împrăştierii datelor. Luarea în considerare a unor astfel de indicatori măreşte valoarea de predicţie a indicatorilor de tendinţă centrală (de măsură a localizării).

Indicatorii de măsură ai împrăştierii trebuie să satisfacă unele condiţii similare celor enunţate pentru indicatorii de măsură a localizării:

- să se bazeze pe toate observaţiile;
- să fie uşor de calculat;
- să fie afectaţi cât mai puţin posibil de fluctuaţiile de selecţie şi să fie adecvaţi pentru un studiu algebric.

Observaţie. Dacă se utilizează un program de calculator care permite obţinerea statisticilor descriptive pentru un set de observaţii, se poate întâmpla ca rezultatele obţinute să difere, într-o oarecare măsură, de rezultatele care s-ar obţine aplicând formulele teoretice, prezentate în continuare. Acest fapt se datorează atât posibilităţii ca, în program, să se utilizeze formule uşor modificate, cât şi unor aproximări inerente procesului de calcul. Recomandarea uzuală este ca, în acest caz, să se utilizeze rezultatele oferite de program[23].

3.1. Amplitudinea

Amplitudinea (sau **domeniul**), notată cu R, este cea mai simplă măsură a împrăştierii. Amplitudinea se defineşte ca fiind drept diferenţa dintre cea mai mare şi cea mai mică valoare observată.

$$R = X_{max} - X_{min}$$

Obiecţiile cu privire la folosirea amplitudinii apar în domenii în care viteza de calcul şi simplitatea (a nu se confunda cu simplismul) în interpretare nu sunt cerinţe importante.

Amplitudinea este afectată de fluctuaţiile de selecţie mari de la eşantion la eşantion şi este de aşteptat ca mărirea numărului de observaţii dintr-un eşantion să ducă la mărirea amplitudinii. Să luăm exemplul timpului de rezolvare (exprimat în zecimi de secundă) a unei dificultăţi:

55, 56, 65, 54, 66, 58, 62, 61, 12, 56, 68, 67, 59, 57, 60, 145.

În acest caz, avem două valori extreme foarte puternic distanţate de restul valorilor, şi anume 12 şi 145 care, probabil, aparţin în primul caz unui supradotat, iar în al doilea caz unui hipofren. Dacă datele de acest gen sunt într-un număr foarte mic, se pot elimina practic din operaţiunile de grupare (procedeu aplicat la unele competiţii sportive unde se elimină notele extreme înainte de a se calcula media unui concurent). Dacă la datele anterioare am efectua calculul amplitudinii, am găsi o valoare de câteva ori mai mare decât în cazul în care eliminăm observaţiile extreme. În primul caz amplitudinea este $X_{max} - X_{min} = 145 - 12 = 133$. În al doilea caz avem: $R = X_{max} - X_{min} = 68 - 54 = 13$. În mod clar, această amplitudine este mai realistă, caracterizând majoritatea datelor. Eliminarea unor valori se va efectua totuşi cu discernământ, deoarece rareori există limite reale superioare sau inferioare pe care valorile unei variabile să nu le poată depăşi. Problema este încadrată, din punct de vedere statistic, în procesarea valorilor aberante.

23. D.S. Moore; G.P. McCabe, *Introduction to the practice of statistics*, ediţia a III-a, W.H. Freeman and Company, New York, 1998, p. 46.

Amplitudinea nu ţine cont de forma repartiţiei; aceeaşi valoare pentru amplitudine se poate obţine atât în cazul unei curbe de frecvenţe simetrice, cât şi în acela al unei curbe de frecvenţe în formă de J. Acelaşi lucru se poate spune despre împrăştierea datelor între valorile extreme.

3.2. Quantile

După definiţia introdusă la prezentarea scalelor ordinale, numim quantilă de ordin p, notată q_p, acea valoare care depăşeşte o proporţie p de observaţii şi este depăşită de o proporţie $1 - p$ de observaţii, unde p este un număr cuprins strict între 0 şi 1.

Pentru date negrupate, calculul unei quantile se realizează după un algoritm similar celui prezentat la datele ordinale:

- se ordonează valorile observate, obţinându-se şirul $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$, unde n este numărul total de observaţii;
- se calculează indicele (rangul) valorii care efectuează împărţirea efectivului de observaţii în proporţiile dorite prin:

$$k = [p \times n] + 1;$$

- quantila de ordin p este atunci valoarea ordonată de rang k, $q_p = x_{(k)}$.

Reamintim că mediana este quantila de ordin 0,5.

Pentru date grupate, se va utiliza o formulă de interpolare similară formulei de calcul al medianei pentru date grupate de interval:

- se calculează indicele k al observaţiei care dă quantila, după formula de mai sus;
- se identifică intervalul căruia îi aparţine observaţia $x_{(k)}$;
- se aplică formula:

$$q_p = l_i + l \times \frac{k - f_{i-1}\downarrow}{f_i},$$

unde: l_i este limita inferioară exactă a intervalului, l este lungimea intervalului, $f_{i-1}\downarrow$ este frecvenţa cumulată pentru intervalul anterior, f_i este frecvenţa (necumulată) a intervalului care conţine quantila.

Şi în cazul datelor de interval se poate vorbi despre sisteme de quantile: quartile, decile, centile etc., după clasificarea prezentată la datele ordinale. Reamintim că se notează prima quartilă cu Q_1, a doua quartilă cu Q_2 şi a treia quartilă cu Q_3.

Utilizarea unui sistem de quantile cu mai multe elemente (decile, centile) conduce la o înţelegere mai fină a modului în care sunt împrăştiate valorile observate în domeniul de variaţie.

3.2.1. Abaterea quartilă

Pornind de la un sistem de quantile, se definesc mărimi derivate, care ajută la caracterizarea gradului de împrăştiere a valorilor observate. Se utilizează astfel:

- **intervalul interquartilic**: IQR = $Q_3 - Q_1$, unde Q_1 şi Q_3 desemnează quartilele 1 şi 3 ale distribuţiei. Denumirea de IQR provine din engleză (InterQuartile Range). În acest interval se află 50% dintre valorile distribuţiei. Cu cât intervalul interquartilic este mai mare, cu atât valorile sunt mai împrăştiate. Este de remarcat faptul că valorile extreme (eventual valori aberante) sunt excluse din calculul acestui indicator prin însăşi formula utilizată. În acelaşi timp, intervalul interquartilic este influenţat de valorile observate doar indirect, prin localizarea acestor valori, şi nu prin valorile efective: două serii cu acelaşi interval interquartilic pot să difere semnificativ ca distribuţie a valorilor;
- **intervalul interdecilic**: $Q_9 - Q_1$, unde Q_9 şi Q_1 sunt decilele 1 şi 9 ale distribuţiei. În acest interval se găsesc 90% dintre valorile distribuţiei. Toate observaţiile de la intervalul interquartilic sunt valabile şi aici.

Se mai poate utiliza **intervalul intercentilic**, marcat prin $Q_{99} - Q_1$, unde prin Q_1 şi Q_{99} notăm centilele 1 şi 99.

Deoarece quartilele nu sunt localizate simetric în domeniul de variaţie a datelor, se introduce conceptul de **abatere quartilă**, prin care se realizează dezideratul de a determina un interval centrat pe mediană (care este quartila a doua), interval în care sunt incluse aproximativ 50% dintre observaţii.

Abaterea quartilă (sau **amplitudinea semiinterquartilă**), notată AQ, este definită drept semidistanţa dintre quartila întâi şi a treia:

$$AQ = \frac{Q_3 - Q_1}{2}.$$

În intervalul cu limitele Me – AQ şi Me + AQ, simetric faţă de mediana Me, se va considera că există jumătate dintre observaţii. Această afirmaţie este adevărată doar în cazul repartiţiilor simetrice (cum este repartiţia normală) şi erorile sunt majore pentru repartiţii cu asimetrie pronunţată. Lungimea intervalului definit este egală cu cea a intervalului interquartilic.

Este de menţionat faptul că, fără o justificare fundamentată teoretic, noţiunile de interval interquantilic şi abatere quartilă sunt uneori utilizate şi pentru date ordinale.

Pentru o distribuţie normală se poate considera că amplitudinea este de aproximativ 7,5 abateri quartile.

Se poate efectua o gradare pe curbă folosind mediana şi abaterea quartilă, la fel cum se procedează în gradarea cu media şi abaterea standard.

3.2.2. Criteriul 1,5×IQR

Pentru identificarea valorilor care par a fi aberante, neconforme cu distribuţia de valori observate, se utilizează frecvent următorul criteriu:

Este aberantă orice observaţie situată la o distanţă mai mare decât 1,5×IQR sub prima quartilă sau peste a treia quartilă.

3.2.3. Rezumarea prin cinci valori

O caracterizare rezumativă a distribuţiei se obţine prin enumerarea următoarelor statistici, calculate din datele observate, în ordinea următoare: valoarea minimă, prima quartilă, mediana, a treia quartilă, valoarea maximă, adică

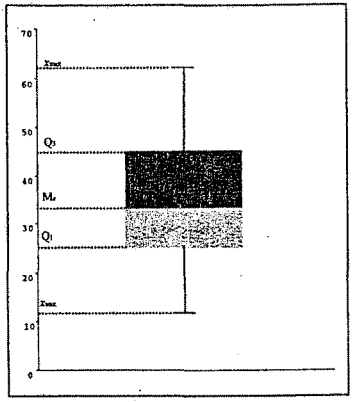
$$x_{min} \quad Q_1 \quad Me \quad Q_3 \quad x_{max}$$

Prin raportarea acestor valori, se obţine o imagine sintetică a amplitudinii datelor (prin includerea valorilor extreme), a tendinţei centrale (prin mediană) şi a împrăştierii datelor (prin quartile). Utilizând valorile enumerate, pot fi calculate şi alte statistici, cum ar fi intervalul interquartilic, abaterea quartilă etc.

Cele cinci valori rezumative se pot vizualiza prin intermediul unei diagrame de tip boxplot. Astfel, o distribuţie caracterizată de valorile

$x_{min} = 12$ $Q_1 = 25$ $Me = 33$ $Q_3 = 45$ $x_{max} = 62$

se reprezintă sub forma



în care sunt prezentate foarte sugestiv mai ales poziţia intervalului interquartilic în domeniul datelor, mărimea intervalului ocupat de fiecare quartilă etc.

3.3. Abaterea medie absolută

Abaterea medie absolută (sau **abaterea medie**), notată cu *AM*, se defineşte ca fiind media aritmetică a abaterilor absolute de la medie:

$$AM = \frac{1}{N}\sum_{i=1}^{N}\left|X_i - \overline{X}\right|.$$

În cazul datelor grupate în k intervale:

$$AM = \frac{1}{N}\sum_{i=1}^{k} n_i \left| c_i - \overline{X} \right|,$$

unde n_i, c_i desemnează frecvența clasei i și, respectiv, centrul clasei de grupare; N reprezintă volumul eșantionului ($N = \sum_{i=1}^{k} n_i$).

În formula de calcul a abaterii medii s-a luat valoarea absolută pentru a se considera mărimea distanței până la medie (fără semn, se știe, de altfel, că suma abaterilor de la medie este egală cu zero). Formulele de calcul arată că AM se bazează pe toate valorile observate și se exprimă în aceleași unități de măsură ca datele observate.

Deoarece toate valorile sunt egal ponderate (vezi puterea întâi a termenilor sumei), se poate obiecta că în aprecierea gradului de împrăștiere ar trebui ca valorile extreme (care definesc oarecum împrăștierea; vezi amplitudinea) să aibă o influență mai mare.

Pentru distribuțiile simetrice se poate considera că în intervalul medie plus/minus o abatere medie se găsesc aproximativ 57% dintre subiecții seriei respective. O abatere standard are aproximativ 1,25 abateri medii.

3.4. Dispersia

Dispersia (sau **varianța**), notată cu s^2, este indicatorul statistic cel mai utilizat pentru aprecierea împrăștierii datelor (atât din rațiuni practice, cât și teoretice). Dispersia se definește ca fiind media aritmetică a pătratelor abaterilor de la medie:

(1) $$s^2 = \frac{1}{N}\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^2.$$

Datorită ridicării abaterilor de la medie la puterea a doua, se realizează ponderarea suplimentară a valorilor extreme (mai depărtate de medie), ceea ce concordă mai bine cu simțul comun în privința aprecierii gradului de împrăștiere. De exemplu, o valoare depărtată de medie cu 10 contribuie la calculul dispersiei cu 100, în timp ce o valoare mai apropiată de medie, având o abatere de 2 de la medie, contribuie doar cu 4. Se observă astfel că „greutatea" contribuției la dispersie este amplificată pentru valorile mai depărtate de medie.

Trebuie precizat faptul că, din punct de vedere teoretic, dispersia reprezintă momentul centrat de ordinul al doilea, fiind deci încadrată într-o clasă de indicatori cu proprietăți remarcabile.

Menționăm că suma de pătrate din definiția dispersiei are proprietatea că este minimă, în sensul că, oricare ar fi o altă valoare de referință a abaterilor, W, are loc:

$$\sum (X_i - W)^2 \geq \sum (X_i - \overline{X})^2,$$

relație care, pe de o parte, aduce un argument suplimentar în favoarea utilizării dispersiei și, pe de altă parte, atribuie mediei calitatea de a fi cea mai bună aproximare, în sensul celor mai mici pătrate, a seriei de date (reamintim că media realiza și minimumul sumei de abateri).

În cazul datelor grupate în k intervale:

(2) $$s^2 = \frac{1}{N}\sum_{i=1}^{k} n_i \left(c_i - \overline{X}\right)^2,$$

unde n_i, c_i desemnează frecvența clasei i și, respectiv, centrul clasei de grupare, iar N este volumul eșantionului ($N = \sum_{i=1}^{k} n_i$).

Deoarece:

$$\sum_{i=1}^{N}(X_i - \overline{X})^2 = \sum_{i=1}^{N} X_i^2 - N\overline{X}^2,$$

formulele de calcul ale dispersiei se pot transforma în

(3) $$s^2 = \frac{1}{N}\left[\sum_{i=1}^{N} X_i^2 - \frac{1}{N}\left(\sum_{i=1}^{N} X_i\right)^2\right]$$

și, respectiv,

(4) $$s^2 = \frac{1}{N}\left[\sum_{i=1}^{N} n_i c_i^2 - \frac{1}{N}\left(\sum_{i=1}^{N} n_i c_i\right)^2\right],$$

unde notațiile sunt cele uzuale.

Notând, pentru ușurința expunerii, dispersia unei variabile X prin $D^2(X)$, se pot demonstra imediat (doar din aplicarea definiției) următoarele proprietăți:

a) Dispersia unei constante este egală cu zero:

$$D^2(c) = 0.$$

b) Dispersia produsului dintre o constantă și o variabilă:

$$D^2(cX) = c^2 D^2(X).$$

c) Dispersia unei sume finite de variabile:

$$D^2(X_1 + X_2 + \ldots + X_k) = D^2(X_1) + D^2(X_2) + \ldots + D^2(X_k).$$

d) Dispersia sumei dintre o constantă c și o variabilă X (caz particular al relației precedente):

$$D^2(c + X) = D^2(X).$$

e) Dispersia diferenței a două variabile:

$$D^2(X_1 - X_2) = D^2(X_1) + D^2(X_2).$$

Observaţie. Reamintim că prin suma a două variabile nu înţelegem reunirea valorilor acestor variabile (cum ar fi considerarea împreună a notelor obţinute de elevii din două clase; în acest caz, notele fiecărei clase constituie valorile unei variabile de sine stătătoare). Pentru această situaţie se va prezenta ulterior formula adecvată.

Proprietăţile dispersiei au aplicaţii practice în reducerea volumului de calcule sau în facilitarea calculelor prin micşorarea ordinului de mărime a valorilor.

De exemplu, dacă valorile variabilei X sunt: 100, 200, 200, 300, 100, 400 etc., atunci se poate considera că X = 100 X', unde X' are valorile 1, 2, 2, 3, 1, 4 etc. Prin urmare:

$$D^2(X) = 100^2 \times D^2(X') = 10000\, D^2(X').$$

Facilitarea calculelor este evidentă: se va calcula dispersia pentru noua variabilă X' şi rezultatul se va înmulţi cu 10000 pentru a obţine dispersia variabilei iniţiale X. Acest exemplu arată totodată modul în care se reflectă în valoarea dispersiei unele transformări aduse valorilor variabilei avute în vedere.

Din considerente teoretice, în cazul utilizării dispersiei de sondaj (calculată dintr-un eşantion) ca estimaţie a dispersiei întregii populaţii statistice (populaţia reprezentată de eşantionul disponibil), formulele (1), (2), (3) şi (4) sunt „corectate" sub forma:

(5) $$s^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^2$$

(6) $$s^2 = \frac{1}{N-1}\sum_{i=1}^{k} n_i\left(c_i - \overline{X}\right)^2$$

(7) $$s^2 = \frac{1}{N-1}\left[\sum_{i=1}^{N} X_i^2 - \frac{1}{N}\left(\sum_{i=1}^{N} X_i\right)^2\right]$$

(8) $$s^2 = \frac{1}{N-1}\left[\sum_{i=1}^{N} n_i c_i^2 - \frac{1}{N}\left(\sum_{i=1}^{N} n_i c_i\right)^2\right],$$

diferenţa faţă de formulele iniţiale fiind împărţirea la N – 1, şi nu la N. În cazul eşantioanelor mari (N > 30), diferenţele dintre rezultate sunt neglijabile.

Observaţie. Deşi notaţiile nu sunt distincte pentru rezultatele celor două grupe de formule, se va preciza explicit numărul formulei atunci când nu va reieşi din context care formulă este utilizată.

Interpretarea practică (intuitivă) a dispersiei este îngreunată pentru că nu se exprimă în aceleaşi unităţi de măsură ca datele iniţiale. De exemplu, dacă se calculează dispersia pentru înălţimi exprimate în cm, atunci dispersia se exprimă în cm^2.

Acest neajuns este înlăturat de următorul indicator statistic, abaterea standard, utilizat frecvent în raportări în locul dispersiei.

3.5. Abaterea standard

Abaterea standard (sau **abaterea medie pătratică**), notată cu *s*, se defineşte ca fiind rădăcina pătrată din dispersie:

$$s = \sqrt{s^2}.$$

Formula trebuie considerată din punct de vedere dinamic, adică se calculează mai întâi dispersia şi apoi abaterea standard, şi nu ca o simplă formulă matematică evidentă.

La fel ca dispersia, din care provine, abaterea standard poate fi considerată un caz particular al unei cantităţi mai generale cunoscute sub numele de abatere medie pătratică de la o valoare arbitrară. Fie A o valoare arbitrară a lui X şi u abaterea lui X de la A, adică u = X – A. Calculele pentru obţinerea abaterilor medii pătratice de la medie şi de la A sunt organizate în tabelul următor:

Nr.crt.	X	x	x^2	u	u^2
1	12	–5	25	–4	16
2	13	–4	16	–3	9
3	14	–3	9	–2	4
4	15	–2	4	–1	1
5	16	–1	1	0	0
6	17	0	0	1	1
7	18	1	1	2	4
8	19	2	4	3	9
9	20	3	9	4	16
10	21	4	16	5	25
11	22	5	25	6	36
TOTAL	187	0	110	11	121

Se obţine:

$$\overline{X} = 17;\ \sum\left(X - \overline{X}\right) = \sum x = 0;\ \sum(X - A) = \sum u = 11;\ \sum\left(X - \overline{X}\right)^2 = \sum x^2 = 110;$$
$$\sum(X - A) = \sum u^2 = 121.$$

Din (1) rezultă $s_1 = 3{,}1623$; din (2) rezultă $s_2 = 3{,}3166$. Utilizând abaterile de la A, obţinem $s_3 = 3{,}3166$. Se observă că aplicarea formulei (1) conduce la un rezultat sensibil mai mic datorită eşantionului de volum mic. Se mai poate nota că suma pătratelor abaterilor de la A este mai mare decât suma pătratelor abaterilor de la medie. Egalitatea dintre s_2 şi s_3 este o pură întâmplare.

3.6. Reuniunea seriilor de date

Dacă se consideră o serie de date compusă din două serii disjuncte, se pot calcula media şi dispersia seriei totale cu ajutorul mediilor şi dispersiilor seriilor componente. Notăm cu N, M şi s numărul de observaţii, media şi abaterea standard a seriei totale, cu N_i, M_i şi s_i, i =1, 2, numărul de observaţii, media şi, respectiv, abaterea standard pentru seriile componente. De notat că $N = N_1 + N_2$.

Din

$$X_{1t} - M = (X_{1t} - M_t) + (M_t - M) = (X_{1t} - M_t) + d_1$$
$$X_{2t} - M = (X_{2t} - M_t) + (M_t - M) = (X_{2t} - M_t) + d_2,$$

unde prin d_1, d_2 se notează diferenţele dintre mediile celor două serii şi media totală. Rezultă, după un calcul algebric elementar, că:

$$(X_{1t} - M)^2 = (X_{1t} - M_t)^2 + d_1^2$$
$$(X_{2t} - M)^2 = (X_{2t} - M_t)^2 + d_2^2,$$

de unde, însumând şi utilizând (1), se obţine relaţia finală:

$$s^2 = \frac{N_1(s_1^2 + d_1^2) + N_2(s_2^2 + d_2^2)}{N}. \quad (9)$$

În cazul în care mediile celor două serii componente sunt egale, formula devine media ponderată a dispersiilor seriilor componente. Dacă, în plus, cele două serii au acelaşi număr de observaţii, atunci dispersia seriei totale se obţine ca fiind media aritmetică a dispersiilor individuale.

Formula (9) poate fi generalizată la cazul a k serii:

$$s^2 = \frac{\sum_{i=1}^{k} N_i(s_i^2 + d_i^2)}{\sum_{i=1}^{k} N_i}.$$

Observaţie. De notat că formula (9) sau generalizarea ei se aplică doar atunci când o populaţie este compusă din subpopulaţii disjuncte (fără elemente comune) şi când se cunosc, cu excepţia dispersiei totale, toţi indicatorii statistici implicaţi. Această situaţie se întâlneşte suficient de rar în practică.

O situaţie deosebită apare atunci când dispunem de indicatorii statistici calculaţi din k eşantioane extrase dintr-o aceeaşi populaţie normală (fără a fi epuizate, evident, toate elementele populaţiei şi fără a fi necesar ca eşantioanele să fie distincte). Estimarea dispersiei totale se face, în acest caz, prin:

$$s^2 = \frac{\sum_{i=1}^{k}(N_i - 1) \times s_i^2}{\sum_{i=1}^{k}(N_i - 1)}, \quad (10)$$

unde dispersiile s_i^2 sunt calculate cu una dintre formulele (5)-(8).

Să luăm două serii de date prezentate în tabelul următor:

Nr. crt	X_1	X_1-M_1	$(X_1-M_1)^2$	X_2	X_2-M_2	$(X_2-M_2)^2$
1	9	–3	9	15	–4,5	20,25
2	11	–1	1	17	–2,5	6,25
3	12	0	0	18	–1,5	2,25
4	10	–2	4	22	2,5	6,25
5	13	1	1	20	0,5	0,25
6	12	0	0	19	–0,5	0,25
7	14	2	4	22	2,5	6,25
8	15	3	9	23	3,5	12,25
Total	96		30	156		254,00

Pentru prima serie obţinem $s_1^2 = \frac{30}{8} = 3{,}75$. Pentru seria a doua $s_2^2 = 31{,}75$. Cum seriile au acelaşi număr de elemente, dispersia întregii serii o constituie media aritmetică a dispersiilor componentelor ei $s^2 = \frac{3{,}75 + 31{,}75}{2} = 17{,}75$.

3.7. Proprietăţi ale abaterii standard legate de curba normală

Amplitudinea unei curbe normale Gauss poate fi aproximată satisfăcător de R = 6 abateri standard. Aplicaţii foarte frecvente privesc calcularea ariilor unor suprafeţe remarcabile de sub curba normală. Se ştie că asemenea arii servesc la

estimarea unor probabilităţi (exprimate uneori sub formă procentuală) de apariţie a valorilor variabilelor ce sunt repartizate normal (o înţelegere mai bună a rezultatelor poate fi obţinută după parcurgerea capitolului dedicat introducerii în teoria probabilităţilor). Astfel, între $\overline{X}-s$ şi $\overline{X}+s$ se află aproximativ 68% din suprafaţa de sub curbă. Între $\overline{X}-2s$ şi $\overline{X}+2s$ se află aproximativ 95% din suprafaţa de sub curbă, iar între $\overline{X}-3s$ şi $\overline{X}+3s$ se află aproximativ 99,958% din suprafaţa de sub curbă, adică practic toată suprafaţa (de aici şi afirmaţia legată de amplitudine).

Această proprietate a abaterii standard constituie baza de pornire pentru a efectua operaţiunea de gradare pe curbă. Alcătuirea de bareme psihologice pe baza unităţilor sigma este o operaţiune frecventă în psihologia aplicată, mai ales în selecţia şi orientarea profesională.

3.8. Construcţia baremelor bazate pe medie şi abatere standard

Să presupunem că avem o distribuţie cu media de $\overline{X}$ şi abaterea standard σ. Pe baza ipotezei că distribuţia variabilei este aproximată suficient de bine de distribuţia normală, rezultă că amplitudinea este de aproximativ 6 abateri standard, adică R = 6σ. După ce am calculat amplitudinea, următorul pas este calcularea bazei de gradare (Bg) în funcţie de numărul de clase în care dorim să facem gradarea:

$$\text{Bg} = R/k,$$

unde k este numărul de clase.

Clasele sunt construite simetric în jurul mediei, au amplitudinea egală cu baza de gradare, iar limitele lor sunt dependente de paritatea numărului de clase. În funcţie de semnificaţia variabilei studiate şi de necesităţile de interpretare, este posibil să se numeroteze clasele în ordine crescătoare (de la valorile mici ale limitelor spre cele mari) sau descrescătoare (de la valorile mari ale limitelor spre cele mici). În tabelele care urmează sunt prezentate, din acest motiv, două coloane cu numerotarea claselor.

a) Număr impar de clase ($k = 2n+1$). Intervalul central, cel cu indicele n + 1, este centrat pe medie; celelalte se obţin din acesta prin scăderea (spre stânga) sau adunarea (spre dreapta) de multipli ai bazei de gradare. Se obţin astfel următoarele intervale de clase:

Clasa	Limita inferioară	Limita superioară	Clasa
1	$\overline{X}-\text{Bg}/2-n\times\text{Bg}$	$\overline{X}-\text{Bg}/2-(n-1)\times\text{Bg}$	2n + 1
2	$\overline{X}-\text{Bg}/2-(n-1)\times\text{Bg}$	$\overline{X}-\text{Bg}/2-(n-2)\times\text{Bg}$	2n
...	...	...	...
n	$\overline{X}-\text{Bg}/2-\text{Bg}$	$\overline{X}-\text{Bg}/2$	n + 2
n + 1	$\overline{X}-\text{Bg}/2$	$\overline{X}+\text{Bg}/2$	n + 1
n + 2	$\overline{X}+\text{Bg}/2$	$\overline{X}+\text{Bg}/2+\text{Bg}$	n
...	...	...	...
2n	$\overline{X}+\text{Bg}/2+(n-2)\times\text{Bg}$	$\overline{X}+\text{Bg}/2+(n-1)\times\text{Bg}$	2
2n + 1	$\overline{X}+\text{Bg}/2+(n-1)\times\text{Bg}$	$\overline{X}+\text{Bg}/2+n\times\text{Bg}$	1

Se observă că a (n + 1)-a clasă este clasa centrală şi că de la limitele ei se poate pleca în ambele sensuri pentru a obţine celelalte limite de clase.

b) Număr par de clase ($k = 2n$). Această situaţie, deşi mai puţin uzuală, nu mai evidenţiază un interval central. Limitele claselor se obţin plecând de la medie şi construind spre stânga şi spre dreapta intervale de amplitudinea Bg:

Clasa	Limita inferioară	Limita superioară	Clasa
1	$\overline{X}-n\times\text{Bg}$	$\overline{X}-(n-1)\times\text{Bg}$	2n
2	$\overline{X}-(n-1)\times\text{Bg}$	$\overline{X}-(n-2)\times\text{Bg}$	2n – 1
...	...	...	...
n	$\overline{X}-\text{Bg}$	$\overline{X}$	n + 1
n + 1	$\overline{X}$	$\overline{X}+\text{Bg}$	n
...	...	...	...
2n – 1	$\overline{X}+(n-2)\times\text{Bg}$	$\overline{X}+(n-1)\times\text{Bg}$	2
2n	$\overline{X}+(n-1)\times\text{Bg}$	$\overline{X}+n\times\text{Bg}$	1

Exemplul 3.1. Fie o distribuţie cu media de 42 şi abaterea standard egală cu 9. Considerând că distribuţia variabilei este aproximată suficient de bine de distribuţia normală, rezultă că amplitudinea este de aproximativ 6 abateri standard, adică R = 6 × 9 = 54. Dorim să facem o împărţire în 7 clase pe care le vom nota cu A, B, C, D, E, F şi G, ordinea fiind descrescătoare (clasa A este caracterizată de limitele de grupare mai mari). Cum numărul de clase este impar, utilizăm procedura a). Avem Bg = 54 : 7 = 7,70 şi Bg/2 = 3,85. Această valoare o adăugăm şi apoi o scădem din valoarea mediei pentru a delimita valorile clasei centrale, clasa D. Aşadar, 42 – 3,85 = 38,15 şi 42 + 3,85 = 45,85 şi am obţinut limitele clasei centrale D. La limita superioară a clasei vom adăuga o bază de gradare şi vom obţine limita superioară a clasei C : 45,85 + 7,70 = 53,55. La această valoare vom adăuga din nou 7,70 şi vom obţine limita superioară a clasei B care, în mod evident, constituie şi limita inferioară a clasei A. Deci 53,55 + 7,70 = 61,25. Pentru a calcula clasa E scădem din valoarea 38,85 (limita inferioară a clasei D) valoarea

7,70 şi obţinem 38,85 – 7,70 = 31,15. Pentru a afla limita inferioară a clasei F şi limita superioară a clasei G, scădem din 31,15 valoarea 7,70.

Deci 31,15 – 7,70 = 23,45.

Observaţie. Valorile obţinute sunt rotunjite la valorile întregi dacă valorile reale obţinute în cursul unui test sau al unui experiment nu se exprimă cu zecimale. Rotunjirea se realizează după regula uzuală: în cazul în care valoarea zecimalelor depăşeşte 0,50, rotunjim prin adăugare; în cazul în care valoarea acestora este sub 0,50, rotunjim prin scădere; în cazul în care valoarea părţii zecimale este egală cu 0,50, procedăm potrivit situaţiei concrete în care ne aflăm, cu precizarea de rigoare.

Clasele astfel obţinute cu ajutorul mediei şi al abaterii standard sunt:

Clasa	Limitele
A	începând cu 61
B	54-60
C	46-53
D	39-45
E	31-38
F	24-30
G	până la 23

Observăm că toate clasele au limite întregi obţinute prin rotunjire, iar jocul rotunjirilor a făcut ca o parte din ele să aibă amplitudini mai mici cu o unitate.

Clasele conţin proporţii diferenţiate de subiecţi. Pentru a stabili ce cotă procentuală de subiecţi conţine fiecare dintre clase, ne bazăm pe aproximarea prin curba normală şi căutăm într-un tabel al funcţiei normate a lui Laplace, într-un tabel al funcţiei de repartiţie normală standard sau într-un tabel similar. Pentru a utiliza un asemenea tabel, se vor calcula valorile transformate z prin:

$$z = \frac{X - \overline{X}}{\sigma}.$$

În exemplul avut în vedere, pentru a calcula efectivul clasei centrale D, procedăm în felul următor: din limita superioară exactă a clasei, 45,5, scădem valoarea mediei şi împărţim la σ, adică z = (45,5 – 42)/9 = 0,39. Pentru acest z (sau pentru valoarea cea mai apropiată de z), în tabelul funcţiei lui Laplace se citeşte 0,1517. Această valoare interpretată procentual (unele tabele dau, de altfel, valoarea direct în procente) arată că în intervalul dintre medie şi limita superioară a lui D se găsesc aproximativ 15,17% dintre subiecţii examinaţi. Având în vedere faptul că D este simetric, centrat pe medie, întreaga clasă va avea 15,17% + 15,17% = 30,34% din totalul subiecţilor. Pentru a calcula cuantumul procentual al clasei C, calculăm transformata z a limitei exacte superioare z = (53,5 – 42)/9 = 1,28. În dreptul

valorii z = 1,28 găsim 39,97%. Dar acest procent conţine şi cuantumul procentual al jumătăţii clasei D. Deci din 39,97% scădem 15,17% şi găsim cuantumul procentual al clasei C, adică 24,8%. Prin simetrie clasa E va avea, de asemenea, 24,8%. Un calcul analog pentru clasa B: z = (60,5 – 42)/9 = 2,05; funcţia Laplace este 47,98%. Dar acest cuantum cuprinde jumătate din clasa D şi clasa C în întregime. Deci din valoarea 47,98% va trebui să scădem 39,97%, cât reprezintă porţiunile respective. Se obţine pentru cuantumul procentual al clasei B valoarea de 8,01%. Prin simetrie, clasa F va avea, de asemenea, 8,01%. Pentru a calcula valoarea procentuală a clasei A, trebuie să scădem din 50%, cât reprezintă jumătatea unei distribuţii simetrice, valoarea 47,98% cât reprezintă totalul cuantumurilor claselor anterioare (până la medie). Deci clasa A va avea 50% – 47,98% = 2,02% şi, prin simetrie, clasa G tot 2,02%. Sintetizând, avem următoarele valori:

Clasa	Limite	Procentaj
A	de la 61	2,02%
B	54-60	8,01%
C	46-53	24,80%
D	39-45	30,34%
E	31-38	24,80%
F	24-30	8,01%
G	până la 23	2,02%

Dacă se presupune că cercetarea s-a efectuat pe 300 de subiecţi, numărul de subiecţi aparţinând unei clase se obţine înmulţind procentajul clasei respective cu numărul total de subiecţi. Astfel, clasa A conţine 300×2,02% = 6,06, adică 6 subiecţi (rotunjirea este obligatorie). În mod analog, clasa centrală D conţine 300×30,34% = 91,02, adică 91 de subiecţi.

Observaţie. Jocul aproximărilor poate conduce la situaţia în care însumarea procentelor de pe ultima coloană nu dă exact 100%.
Pentru o discuţie completă a definirii şi utilizării curbei normale se va studia şi secţiunea dedicată acestei repartiţii teoretice în capitolul „Elemente de teoria probabilităţilor".

3.9. Transformări de variabile

3.9.1. Variabile normate

Abaterea standard este utilizată şi pentru obţinerea variabilelor normate. Aceste variabile îşi datoresc apariţia nevoii de obţinere a unor valori standard individualizate. În lipsa unor astfel de valori, transformarea notelor brute în note standard se face prin asumarea codului unei clase valorice.

Cele mai cunoscute variabile normate sunt cele notate cu z, T şi H.

Variabilele z se obţin prin formula uzuală $z = (X - \overline{X})/s$. Efectuând această transformare, se obţin uzual valori între – 3 şi + 3, deşi teoretic se poate obţine orice altă valoare (oricât de mare sau oricât de mică).

Dacă transformăm toate notele unei distribuţii în valori normate z, obţinem o nouă distribuţie pentru care media este egală cu 0, iar abaterea standard este 1. Dificultatea folosirii unor astfel de note constă în faptul că se pot obţine valori negative (mai greu de interpretat în practică) şi, de asemenea, valori cu zecimale. Având în vedere faptul că distribuţia este foarte îngustă, nu sunt recomandate rotunjirile.

Variabilele T suplinesc în bună parte astfel de dificultăţi. Aceste variabile au fost create de americanul McCall la începutul secolului XX. Ele se obţin prin formula:

$$T = 50 + \frac{10}{s}\left(X - \overline{X}\right).$$

Având în vedere definiţia transformării z, putem spune că valorile T înseamnă 50 + 10z. Dacă transformăm datele în valori T, obţinem o nouă distribuţie în care media este egală cu 50 şi abaterea standard cu 10 (demonstraţie imediată din proprietăţile dispersiei). Aceste valori oscilează uzual între 15 şi 85.

Pentru a crea un sistem de valori care să fie în domeniul 0-100, Hull a propus o formulă care a fost preluată de mulţi practicieni în psihologie:

$$H = 50 + \frac{14}{s}\left(X - \overline{X}\right).$$

Din definiţia transformării z rezultă H = 50 + 14z.

3.9.2. Transformări liniare

Toate transformările prezentate în secţiunea anterioară se încadrează în clasa transformărilor liniare, adică a celor de forma x2= ax1 + b, în care x1 este măsura iniţială şi x2 este măsura transformată. Coeficienţii a şi b ai transformării sunt daţi de formulele:

$$a = \frac{\sigma_2}{\sigma_1} \text{ şi } b = m_2 - am_1,$$

unde m_1 şi σ_1 reprezintă media şi, respectiv, abaterea standard a distribuţiei măsurilor iniţiale, iar m_2 şi σ_2 valorile corespunzătoare pentru distribuţia mărimii transformate. Pentru transformarea în scoruri centrate reduse sau scoruri z, avem

$m_2 = 0$, $\sigma_2 = 1$ şi, prin urmare, $a = \frac{1}{\sigma_1}$, $b = -\frac{m_1}{\sigma_1}$. Formula de transformare devine: z $= (x_1 - m_1)/\sigma_1$; cu alte cuvinte, regăsim formula prezentată anterior pentru transformata z.

Tabelul următor conţine o prezentare sintetică a principalelor transformări utilizate în practica de conversiune a scorurilor. Pentru fiecare transformare se precizează media şi abaterea standard a variabilei transformate (care, după formulele anterioare, permit calculul coeficienţilor transformării), precum şi câteva exemple de valori transformate.

Transformare	m	σ	Corespondenţe între diferite scoruri transformate liniar		
Scor original	37,4	6,5	19	25	43
Scoruri z	0,00	1	–2,83	–1,91	0,86
Scoruri Stanford-Binet	100	16	54,71	64,48	113,78
Scoruri Wechsler	100	15	57,54	74,38	112,92
Scoruri H	50	14	10,37	23,29	62,06
Army General Classification Test	100	20	43,38	61,85	117,23
Scala Wechsler (pentru subteste)	10	3	1,51	4,28	12,58
College Board Scholar Aptitude	500	100	216,92	309,29	586,15
Scoruri Stanine	5	2	–0,66	1,18	6,72
Graduate Record Examination	500	100	216,92	309,29	586,15
Scoruri STEN	5,5	2	–0,16	1,68	7,22
California Entrance Examination Board	500	100	216,92	309,29	586,15
Scoruri T	50	10	21,89	30,92	58,82
% > scor			99,76	97,20	19,50

Ultima linie a tabelului conţine procentajul de subiecţi care, într-o distribuţie normală, au scoruri superioare scorului original.

Entropia, notată cu H (a nu se confunda cu transformata Hull), a fost introdusă în 1948, de C. Shanon, în teoria statistică a informaţiei. Dintre posibilele sensuri ale noţiunii, în analiza datelor este utilizat acela care priveşte entropia ca o măsură a gradului de dezordine a unui sistem. Astfel, entropia poate fi folosită ca o măsură a împrăştierii. Formula de calcul[24] este următoarea:

$$H = \sum_{i=1}^{k} p_i \log_2 \frac{1}{p_i} = -\sum_{i=1}^{k} p_i \log_2 p_i,$$

unde k este numărul de clase în care se împart datele, iar p_i probabilitatea ca un element să se afle în clasa i. Este necesar ca toate probabilităţile să fie strict

24. I.P. Vasilescu, *Statistică informatizată pentru ştiinţe despre om*, vol. 1, Editura Militară, Bucureşti, 1991, p. 32.

pozitive (adică diferite de zero). Domeniul de variație este între 0 și $\log_2 k$ și rezultatul se exprimă în biți (bit este prescurtarea de la *binary digit*).

Entropia poate fi calculată și utilizând logaritmi naturali (în baza e = 2,7183...), caz în care rezultatul este exprimat în niți (de la nit, prescurtare pentru *natural digit*). Formula devine:

$$H = \sum_{i=1}^{k} p_i \ln \frac{1}{p_i} = -\sum_{i=1}^{k} p_i \ln p_i .$$

Aplicând noua formulă, se modifică și domeniul de valori, dar este important faptul că variația rămâne limitată. Pentru o distribuție cu concentrare maximă (toate valorile se situează într-o singură clasă, adică împrăștierea datelor este minimă) se obține H = 0. Valoarea entropiei crește o dată cu mărirea împrăștierii datelor, care poate fi interpretată ca o creștere a dezorganizării, valoarea maximă fiind obținută atunci când elementele sunt repartizate uniform în clasele de grupare.

În aplicații, entropia poate avea doar valoare de comparare. Utilitatea ei apare atunci când se compară două clasificări diferite cu același număr de clase, cea cu entropia mai mare având o împrăștiere crescută a datelor.

Un indicator adimensionalizat, care să permită aprecierea unei valori individuale a entropiei, poate fi obținut, după același principiu aplicat la definirea indicelui de variație calitativă (introdus pentru scale nominale), prin raportarea valorii calculate la valoarea maximă a entropiei. Se introduce astfel noțiunea de **entropie relativă**, notată H_r, care în varianta exprimării în niți este definită prin:

$$H_r = \frac{H}{\ln k} .$$

Exemplul 3.2. Considerăm că în domeniul analizei distractorilor se poate aplica respectivul calcul al entropiei pentru a analiza calitatea construcției itemilor.

Distractorul este folosit doar în cazul testelor care beneficiază de un dispozitiv de răspuns cu alegere multiplă (*multiple choices*). Dintr-un număr de variante propuse ca posibil corecte, subiectul trebuie să aleagă numai una. Dintre variantele propuse, un singur răspuns este corect, restul variantelor constituind așa-numiții distractori. Inadecvările itemilor în ansamblul testului își pot avea sursa în dificultatea constructorilor de teste de a găsi distractori verosimili. Dacă un test ce prezintă un dispozitiv de răspuns cu alegere multiplă nu are distractori verosimili, cel care efectuează testul poate descoperi foarte ușor răspunsul corect. În cazul în care la un test de cunoștințe punem întrebarea: „Cine a scris *Fefeleaga*?" și se indică printre posibilele răspunsuri Ion Agârbiceanu, William Shakespeare, Platon, Dante, Cervantes, cu siguranță că subiectul va „ști" cine a scris nuvela respectivă, deoarece distractorii propuși sunt total neverosimili. Răspunsurile s-ar concentra asupra variantei corecte într-o mare proporție, ceea ce ne-ar putea duce la decizia pripită de a schimba itemul respectiv, deoarece nu prezintă nici o dificultate. Dar, dacă dispozitivul de răspuns ar cuprinde distractori cu numele unor scriitori români

ca Liviu Rebreanu, Ion Slavici, Ion Agârbiceanu și Gala Galaction, cu siguranță vom crea mai multe probleme unui subiect nesigur și vom avea opțiuni mult mai diversificate. Distractorul reprezintă un domeniu de cercetare insuficient explorat de către cei care construiesc instrumente psihometrice de investigație[25].

Redăm, în continuare, situația alegerii răspunsurilor la Matricea Progresivă Raven, forma standard, efectuată pe un lot de 151 de muncitori de la RAJAC Iași, examinați în luna august a anului 1998 (vârsta între 30 și 45 ani, studii generale și școală profesională, proveniență din mediul urban). Alegerea unei anumite variante de răspuns este exprimată în valori absolute în următorul tabel:

Item	V1	V2	V3	V4	V5	V6	V7	V8	Total
E6	11	11	6	14	80	5	5	19	151
E7	60	7	12	17	8	25	15	7	151
E8	18	20	5	15	8	67	7	11	151

Se observă că s-au considerat trei itemi, E6, E7, E8, fiecare prezentând 8 variante de răspuns, notate V1-V8. Coloanele cuprind efectivele de subiecți care au optat pentru diferite variante de răspuns. La itemul E6 varianta corectă este V5, la itemul E7 este V1, iar la E8 este V6.

În tabelul următor prezentăm proporția diferitelor variante, rezultată prin împărțirea numărului opțiunilor pentru o anumită variantă la numărul total de răspunsuri (151). Astfel, 0,072 (intersecția liniei E6 cu coloana V1) reprezintă proporția subiecților care au optat pentru V1 la itemul E6. De la stânga la dreapta, pe o linie, aceste frecvențe relative constituie estimații pentru probabilitățile p_1, p_2, $p_3, \ldots, p_8$, implicate în formula de calcul al entropiei:

Item	V1	V2	V3	V4	V5	V6	V7	V8	Total
E6	0,0728	0,0728	0,0397	0,0927	0,5298	0,0331	0,0331	0,1258	1,000
E7	0,3974	0,0464	0,0795	0,1126	0,0530	0,1656	0,0993	0,0464	1,000
E8	0,1192	0,1325	0,0331	0,0993	0,0530	0,4437	0,0464	0,0728	1,000

În continuare, prezentăm valorile logaritmilor naturali ai respectivelor proporții, necesare pentru calcularea entropiei răspunsurilor la fiecare item supus analizei:

Item	V1	V2	V3	V4	V5	V6	V7	V8
E6	–2,619	–2,619	–3,226	–2,378	–0,635	–3,408	–3,408	–2,073
E7	–0,923	–3,071	–2,532	–2,184	–2,938	–1,798	–2,309	–3,071
E8	–2,127	–2,022	–3,408	–2,309	–2,938	–0,813	–3,071	–2,619

25. Vezi argumentația în Paul Dick; Jocelyne Tournois; André Flieller; Jean-Luc Kop, *La Psychométrie*, PUF, Paris, 1994, pp. 87-88.

În căsetele tabelului care urmează sunt trecute produsele dintre proporţiile opţiunilor pentru o anumită variantă de răspuns şi logaritmii naturali ai proporţiilor:

Item	V1	V2	V3	V4	V5	V6	V7	V8	Total
E6	–0,191	–0,191	–0,128	–0,220	–0,337	–0,113	–0,113	–0,261	–1,553
E7	–0,367	–0,142	–0,201	–0,246	–0,156	–0,298	–0,229	–0,142	–1,781
E8	–0,254	–0,268	–0,113	–0,229	–0,156	–0,361	–0,142	–0,191	–1,713

Aplicând formula entropiei, măsurate în niţi, se obţin valorile 1,553 (pentru E6), 1,781 (pentru E7) şi 1,713 (pentru E8). Dacă împărţim fiecare dintre aceste valori la valoarea logaritmului natural din 8 (ln8 = 2,079), obţinem valorile entropiei relative ale repartiţiei răspunsurilor la respectivii itemi. Aceste valori sunt 0,75 (pentru E6), 0,86 (pentru E7) şi 0,82 (pentru E8). Concluzia este că itemul E7 are cea mai bună verosimilitate a itemilor şi, din punctul de vedere al dispozitivului de răspuns, este itemul cel mai bine construit.

Capitolul 4

PREZENTAREA ŞI SISTEMATIZAREA DATELOR

În cazul în care sunt efectuate pentru prima dată şi deci nu găsim elemente de raportare, testele şi experimentele psihologice oferă rezultate sub formă numerică brută, care aduc foarte puţină relevanţă.

De regulă, este acceptat faptul că orice informaţie statistică are ataşate trei valori:

- **valoarea istorică**: indică o stare de fapt existentă la un moment dat;
- **valoarea de comparaţie**: permite ca situaţia evaluată să poată fi comparată cu unele situaţii similare;
- **valoarea de predicţie**: permite extrapolarea rezultatului determinat către situaţii viitoare sau către situaţii similare celei examinate.

Transformările pe care le suportă datele brute în timpul prelucrării statistice au şi scopul de a obţine rezultate pentru care este avantajat un anumit tip de valoare, dintre cele enumerate.

4.1. Proceduri generale

În această secţiune se prezintă transformările primare suportate de datele brute, obţinute imediat după corectarea unui test psihologic, după realizarea unui experiment, fără nici o intervenţie prealabilă.

Exemplul 4.1. Exemplificăm printr-o posibilă enumerare de rezultate obţinute în urma examinării a 64 de subiecţi cu testul Praga. În cazul în care datele provin dintr-o examinare reală, este necesar să precizăm data examinării, vârsta subiecţilor, sexul, nivelul de şcolarizare profesională, adică un număr de informaţii care pot fi considerate surse ale variabilităţii interindividuale. Grila de corecţie conduce la rezultate corecte, greşite sau la omisiuni. În exemplu sunt luate în considerare doar rezultatele corecte. Un subiect poate obţine, în mod teoretic, scoruri între 0 şi 100.

Tabelul 4.1. *Datele absolute pentru exemplul 4.1.*

32	37	45	71	74	52	63	39
49	53	58	27	37	40	57	52
52	65	72	43	54	62	48	54
64	50	52	48	58	55	49	68
35	49	54	58	54	58	51	39
41	46	53	69	54	66	47	49
58	75	37	44	53	59	39	66
47	53	58	63	52	39	47	78

Din această înşiruire de date nu putem trage nici o concluzie dacă nu efectuăm o operaţie de ordonare. Este important de stabilit înaintea unor prelucrări de date dacă variabilele studiate sunt discrete sau continue. Cele continue, măsurate pe o scală de interval, permit inserarea unei valori intermediare între oricare alte două valori. Variabilele discrete, măsurate pe scale nominale sau ordinale, nu au această proprietate. Variabilele discrete sunt des utilizate în psihologie şi valorile lor formează mulţimi finite.

Exemple de variabile continue:

- înălţimea unei persoane;
- timpul de efectuare a unei probe;
- temperatura;
- vârsta.

Exemple de variabile discrete:

- numărul de membri ai unei familii;
- numărul de studenţi la o oră de curs;
- scorul la diferite probe.

Observaţie. Deoarece unele variabile au ca valori posibile observate o mulţime discretă finită, aprecierea caracterului de continuitate poate produce confuzii. Exemplificăm prin sistemul de notare din învăţământul nostru cu note între 1 şi 10. Deşi notele uzuale sunt 1, 2, ..., 10, totuşi variabila trebuie să fie considerată continuă, fapt reflectat, de altfel, în notări mai nuanţate şi în medierile care se utilizează. Aceeaşi situaţie se întâlneşte uneori în cazul variabilelor ordinale care au un număr important de gradaţii şi care pot fi considerate, fără a se produce o mare eroare, variabile continue şi deci beneficiare ale prelucrărilor adecvate acestora.

În cazul exemplului anterior, vom considera că datele provin dintr-o variabilă continuă, dar având valori observabile discrete.

Într-o primă fază, o imagine mai clară a ansamblului de valori obţinute se poate crea prin ordonarea crescătoare a acestora:

Tabelul 4.2. *Datele absolute ordonate pentru exemplul 4.1.*

27	32	35	37	37	37	39	39
39	39	40	41	43	44	45	46
47	47	47	48	48	49	49	49
49	50	51	52	52	52	52	53
53	53	53	54	54	54	54	54
55	57	57	58	58	58	58	58
58	59	61	63	63	64	65	66
66	68	69	71	72	74	75	78

Sub această formă de prezentare putem indica în mod sigur limitele extreme ale împrăştierii numerice şi putem observa „din ochi" anumite repetări ale valorilor.

Primul pas în ordonarea unei distribuţii valorice îl constituie aflarea amplitudinii.

În etologie şi fiziologie, de unde este preluat termenul, amplitudinea are sensul de depărtare între valorile extreme ale unui fenomen înregistrat în mod permanent pe o perioadă dată.

În statistică, amplitudinea (notată R) are sensul de diferenţă între valoarea cea mai mare (X_{max}) şi valoarea cea mai mică (X_{min}) a unei variabile, reprezentând domeniul pe care se întind valorile observate.

În cazul repartiţiei noastre de valori $78 - 27 = 51$. De notat că, datorită modului discret în care sunt observate valorile, nu trebuie să confundăm amplitudinea cu numărul de valori distincte pe care le putem observa. Acest număr se obţine, evident, din $78 - 27 + 1 = 52$ şi are semnificaţia că în domeniul de lungime 51 (= amplitudinea) putem observa 52 de valori **întregi** distincte (teoretic, în acest domeniu am putea observa o infinitate de valori **reale**, în sensul atribuit în matematică acestui cuvânt).

Deşi amplitudinea are un sens foarte clar, nu-i putem acorda o valoare informaţională mare, deoarece sunt evidente o serie de inconveniente:

- amplitudinea nu depinde decât de termenii extremi care constituie excepţii în marea majoritate a studiilor;
- amplitudinea este supusă fluctuaţiilor mari de la un eşantion la altul; eşantioanele mari tind să mărească amplitudinea;
- amplitudinea ignoră forma distribuţiei valorilor intermediare; distribuţii foarte diferite între ele pot să aibă aceeaşi valoare a amplitudinii.

În cazul testului psihologic, se poate vorbi de o *amplitudine de construcţie* şi una *de execuţie*. Cea de construcţie este dată de extremele punctajelor prevăzute de autorii testelor. Exemplificăm amplitudinea de construcţie întâlnită la câteva teste de forma creion – hârtie:

- Bourdon-Amfimov
  (destinaţie psihodiagnostică – atenţie concentrată) 0-762
- A.D. Praga (atenţie distributivă) 0-100

- J.L. Raven (capacitate operaţională a gândirii) 0-60
- Kraepelin (oboseală neuropsihică) 0-540

Practica aplicării testelor psihologice arată că, la aceste instrumente de investigaţie standardizate, valorile extreme nu se realizează niciodată sau se obţin foarte rar. După experienţa noastră, în cele 10 minute prevăzute în protocol pentru efectuarea probei Bourdon-Amfimov, doar 1 subiect din 5000 termină proba şi nu s-a întâmplat niciodată ca un subiect să obţină 0 puncte în cazul în care explicaţia efectuării probei a fost înţeleasă. Valoarea 0 nu ar avea nici un sens. Poate spune cineva ce reprezintă o inteligenţă de nivel 0 sau valoarea 0 a altei aptitudini cognitive? Excludem aserţiunile cu efect retoric în această privinţă. Deci, dacă la proba Bourdon-Amfimov amplitudinea teoretică sau de construcţie este egală cu 762, cea practică sau de execuţie nu atinge această valoare. Situaţia se întâlneşte şi la alte probe psihologice.

Următorul pas în vederea organizării datelor în scopul creşterii gradului lor de relevanţă îl constituie gruparea. Ea poate fi efectuată prin stabilirea unor variante de variaţie sau a unor intervale de variaţie.

Punând în aplicare procedeul stabilirii variantelor de variaţie, observăm mai întâi că există $X_{max} - X_{min} + 1$ valori posibile distincte (în exemplul nostru, 52). În practică nu sunt obţinute ca rezultate toate aceste valori distincte, după cum unele pot să apară în mod repetat. Ordonarea pe variante de variaţie poate consemna şi valorile cu frecvenţa 0, obţinându-se tabelul:

Tabelul 4.3. *Frecvenţe după variantele de variaţie (frecvenţele nule incluse)*

varianta	frecvenţa	varianta	frecvenţa	varianta	frecvenţa	varianta	frecvenţa
27	1	40	1	53	4	66	2
28	0	41	1	54	5	67	0
29	0	42	0	55	1	68	1
30	0	43	1	56	0	69	1
31	0	44	1	57	2	70	0
32	1	45	1	58	6	71	1
33	0	46	1	59	1	72	1
34	0	47	3	60	0	73	0
35	1	48	2	61	0	74	1
36	0	49	4	62	1	75	1
37	3	50	1	63	2	76	0
39	4	52	4	65	1	78	1
38	0	51	1	64	1	77	0

În cazul în care numărul variantelor cu frecvenţe nule este relativ mare, interpretarea tabelului este dificilă şi se pierde astfel avantajul operaţiei de grupare. Din acest motiv, este mai comod şi mai util să enumerăm doar variantele cu frecvenţe nenule, deci numai pe acelea care au apărut efectiv în investigare, obţinând astfel tabelul 4.4.

Tabelul 4.4. *Frecvenţe după variantele de variaţie (frecvenţe nenule)*

varianta	frecvenţa	varianta	frecvenţa	varianta	frecvenţa	varianta	frecvenţa
27	1	44	1	52	4	63	2
32	1	45	1	53	4	68	1
35	1	46	1	54	5	69	1
37	3	47	3	55	1	71	1
39	4	48	2	57	2	72	1
40	1	49	4	58	6	74	1
41	1	50	1	59	1	75	1
43	1	51	1	62	1	78	1

Eliminarea frecvenţelor nule are totuşi dezavantajul că, în cazul unui număr suficient de mare de variante, este greu de sesizat câte variante nu au fost prezente şi care sunt acestea. În general, forma de prezentare a datelor într-un tabel este indicată atunci când se urmăreşte inventarierea unor date (ca bază pentru alte prelucrări) sau facilitarea unor comparaţii (între linii sau între coloane). În cazul tabelului 4.4 perceperea globală a datelor este dificil de realizat. O primă imagine generală poate fi obţinută dacă se utilizează o reprezentare grafică a datelor. Forma cea mai uzuală de reprezentare grafică este histograma frecvenţelor.

Figura 4.1 prezintă histograma obţinută pentru datele din tabelul 4.4:

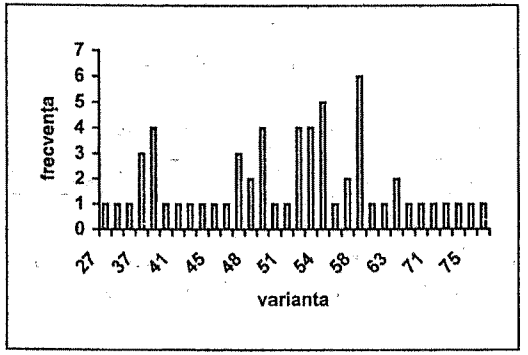


Figura 4.1. *Histograma frecvenţelor*

Prezentate sub această formă, datele pot furniza mărimea amplitudinii şi cele 32 de variante observate din totalul de 52 posibile. Acest gen de grupare este indicat mai ales în cazul existenţei unui număr redus de variante. În histograma anterioară se observă că numărul relativ mare de variante nu asigură o concentrare suficientă a informaţiei şi, de altfel, nu permite, din această cauză, nici desenarea corectă a tuturor detaliilor reprezentării.

Gruparea datelor prin intermediul intervalelor de variaţie este recomandată atunci când caracteristica de grupare prezintă un mare număr de valori individuale. În prelucrările statistice ale cercetării psihologice se practică mai ales gruparea

datelor pe intervale egale, dar nu este exclusă nici cea pe intervale inegale când condiţiile concrete o impun. Intervalele egale sunt de preferat atunci când stabilirea unei mărimi a intervalului nu conduce la obţinerea unui număr prea mare de clase. Intervalele inegale sunt adecvate pentru caracteristici de grupare cu un grad foarte mare şi neuniform de variaţie.

În cazul optării pentru varianta intervalelor, pentru început se stabileşte numărul de intervale de grupare. Deoarece gruparea datelor se impune din necesitatea de concentrare a informaţiei, ar trebui ca numărul intervalelor să fie mic. Deoarece prin grupare se pierde informaţie (nu se mai ştie valoarea exactă observată), ar trebui ca numărul intervalelor să fie mare. Din această contradicţie a criteriilor rezultă că nu există un număr teoretic optim de intervale. Sunt propuse diverse formule, calitatea lor principală fiind doar aceea de uniformizare.

O formulă acceptată de majoritatea statisticienilor este relaţia empirică a lui H.D. Sturges:

$$k = 1 + 3{,}332 log_{10} N,$$

unde k desemnează numărul de intervale, iar N numărul de observaţii (volumul eşantionului).

Lungimea i a unui interval este dată de relaţia:

$$i = \frac{X_{max} - X_{min}}{k}.$$

Utilizând formula lui Sturges, se obţine tabelul 4.5 orientativ pentru stabilirea numărului de grupe:

Tabelul 4.5. *Numărul de clase de grupare (formula lui Sturges)*

Nr. observaţii	15-24	25-44	45-89	90-179	180-359	360-719	720-1439
Nr. grupe	5	6	7	8	9	10	11

Există recomandări diferite în privinţa numărului de grupări sau clase de interval. Pentru valori mici ale lui N există formula lui Kohen-Shapiro:

$$k = \left[\frac{N}{5}\right],$$

unde k este numărul de clase, N volumul eşantionului, iar $[x]$ este partea întreagă a lui x.

Se cunosc, de asemenea, tabelele lui Bendat, Piersol şi Kolker[26].

G. Biji şi N. Biji recomandă ca numărul de intervale să nu depăşească 15. Yulle şi Kendal susţin că numărul optim de clase de interval este cuprins între 15 şi 25.

26. Vezi I.P. Vasilescu, *Statistică informatizată pentru ştiinţe despre om*, vol. 1, Editura Militară, Bucureşti, 1991, p. 27.

Un alt autor, G. Baillargean[27] afirmă că numărul de grupe trebuie să fie cuprins între 15 şi 20.

Dumitru Porojan[28] recomandă formula lui Sturges şi experienţa statistică. În tehnica grupării recomandă respectarea anumitor principii:

- numerele prin care se exprimă intervalele să fie pe cât posibil rotunde;
- numărul de intervale să fie suficient de mare în scopul de a furniza informaţii cât mai analitice;
- intervalele trebuie să permită regruparea datelor sau desfacerea în două;
- frecvenţele mici trebuie să fie evitate, acestea având o semnificaţie statistică redusă[29].

Observăm că în privinţa alegerii numărului de clase de grupare se lasă o libertate destul de mare specialistului care are experienţă în domeniu. În cercetările psihologice, cele mai frecvente grupări sunt în 5, 7, 9, 11 sau 13 clase. De obicei, se alege un număr impar de clase pentru o serie de avantaje ce rezidă în comodităţi de calcul (de exemplu, se poate observa mai bine simetria distribuţiei, există o clasă centrală etc.). Din punct de vedere teoretic însă, nu există restricţii cu privire la alegerea unui număr par de intervale de grupare.

Deoarece operaţia de grupare reprezintă o chestiune importantă în organizarea datelor, este posibil să se apeleze la mai multe variante de grupare până când se ajunge la o soluţie adaptată situaţiei cercetate.

Prin gruparea datelor şi observarea frecvenţelor de interval, valorile individuale sunt mascate şi cea mai bună aproximare a lor este oferită, pe baza principiului erorii minime, de centrul clasei căreia îi aparţin. Este firesc, în această situaţie, să ne aşteptăm la distorsiuni mai mari sau mai mici ale valorilor adevărate. În cazul constatării unui grafic în „J" sau în „L", posibilităţile de deformare a unor indici sunt mult mai mari prin luarea ca valoare reprezentativă a centrului intervalului. În acest caz particular, micşorarea mărimii intervalului constituie calea sigură pentru reducerea distorsiunii (implicit, va creşte numărul de intervale). Analiza erorilor la probele Bourdon-Amfimov sau Kraghin conduce la o astfel de situaţie.

Pentru exemplul 4.1 operaţia de grupare în nouă clase este descrisă în continuare.

Mărimea unui interval este:

$$i = \frac{X_{max} - X_{min}}{9} = \frac{78-27}{9} = \frac{51}{9} = 5{,}7.$$

Atunci când lucrăm cu variabile discrete sau cu variabile pentru care măsurăm doar valori discrete, este indicat să rotunjim la un număr întreg, prin adaos,

27. Vezi Al. Isaic, *Statistica teoretică economică*, Editura Tehnică, Chişinău, 1994, p. 45.
28. Vezi D. Porojan, *Statistica şi teoria sondajelor*, Casa de editură şi presă „Şansa", Bucureşti, 1993, p. 34.
29. Considerăm că în practica cercetării există situaţii în care observaţiile cu frecvenţe reduse (aşa-zisele observaţii aberante) sunt importante prin aceea că surprind caracteristici importante ale fenomenului studiat.

lungimea intervalului calculată anterior. Obţinem astfel o lungime de interval egală cu 6. Vom căuta, în plus, ca fiecărui interval definit să-i aparţină un acelaşi număr de variante de variaţie. Analiza arată că într-un interval de lungime 6 putem avea incluse 6 variante de variaţie întregi, ceea ce produce în total $9 \times 6 = 54$ de valori. Cum situaţia analizată are doar $78 - 27 + 1 = 52$ de variante de variaţie, vom mări, în mod artificial, amplitudinea prin luarea în considerare a încă două variante de variaţie, câte una la fiecare extremitate a domeniului observat. Această extindere poate produce o uşoară disproporţie în cazul în care ar trebui să adăugăm un număr impar de variante. Observăm totuşi că extensia domeniului observat în sensul de mai sus are implicaţii doar asupra frecvenţelor din clasele marginale, care oricum sunt judecate sub aspectul singularităţilor şi nu afectează clasele de mijloc.

Obţinem astfel tabelul 4.6 de frecvenţe, tabel în care sunt indicate clasele prin variantele de variaţie care le caracterizează.

Tabelul 4.6. *Distribuţia frecvenţelor după clasele de grupare*

Clasa		Frecvenţa
interval	cod	
26-31	c1	1
32-37	c2	5
38-43	c3	7
44-49	c4	12
50-55	c5	16
56-61	c6	9
62-67	c7	7
68-73	c8	4
74-79	c9	3

Pentru stabilirea empirică a frecvenţelor, putem utiliza procedeul de numărare popular constând în parcurgerea şirului de valori înregistrate şi trasarea unei liniuţe în dreptul clasei căreia îi aparţine valoarea procesată. Fiecare a cincea linie dintr-o clasă va fi trasată transversal peste precedentele patru pentru o mai uşoară totalizare în final (𝍸). De exemplu, pentru intervalul 50-55 s-ar obţine 𝍸 𝍸 𝍸 |, adică 16. De remarcat că gruparea finală obţinută respectă atât principiul general al intervalelor egale, cât şi avantajul evident al includerii în fiecare clasă a unui număr egal de variante de variaţie.

Prezentarea sub formă de diagramă cu bare a datelor grupate din tabelul 4.6 este redată în figura 4.2. Micşorarea numărului de bare, în raport cu figura 4.1, are ca rezultat evidenţierea mai accentuată a tendinţei centrale, adică a valorilor cel mai des întâlnite în cadrul experimentului.

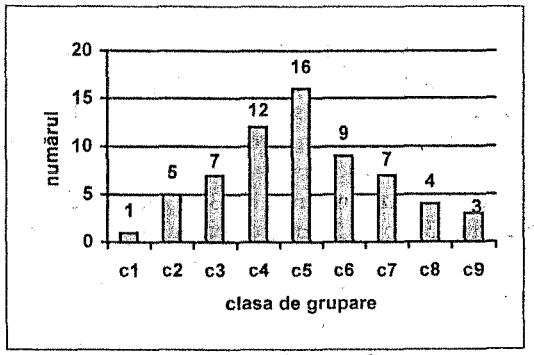


Figura 4.2. *Diagrama frecvenţelor grupate*

Datele, aşa cum se prezintă în tabelul 4.6, sunt înţelese frecvent ca o *serie* sau o *repartiţie statistică*, adică sub forma a două şiruri paralele de numere în care primul conţine enumerarea gradientelor variabilei, iar al doilea frecvenţele corespunzătoare. Acest tabel poate fi completat, obţinându-se tabelul statistic al prelucrării. Tabelul statistic constituie o formă de prezentare raţională a datelor statistice în care valorile observate şi derivatele lor sunt evidenţiate într-un sistem corelat de caracteristici sub formă tabelară.

Pentru exemplul 4.1, tabelul 4.7 exemplifică un tabel statistic conţinând rubrici utilizate frecvent în aplicaţiile psihologice.

Tabelul 4.7. *Tabelul statistic pentru exemplul 4.1.*

Nr. crt.	Intervalul de clasă	Centrul intervalului	f	$\tilde{f}$ (%)	Fc↓	Fc↑	$\tilde{Fc}$ ↓ (%)	$\tilde{Fc}$ ↑ (%)
1	2	3	4	5	6	7	8	9
1	26-31	28,5	1	1,6	1	64	1,6	100,0
2	32-37	34,5	5	7,8	6	63	9,4	98,4
3	38-43	40,5	7	10,9	13	58	20,3	90,6
4	44-49	46,5	12	18,8	25	51	39,1	79,7
5	50-55	52,5	16	25,0	41	39	64,1	60,9
6	56-61	58,5	9	14,1	50	23	78,2	35,9
7	62-67	64,5	7	10,9	57	14	89,1	21,8
8	68-73	70,5	4	6,2	61	7	95,3	10,9
9	74-79	76,5	3	4,7	64	3	100,0	4,7
			N = 64					

Prima coloană conţine numărul de ordine al intervalelor.

A doua rubrică prezintă intervalele în ordine crescătoare. Scrierea în ordine descrescătoare este la fel de practică. Într-un interval are importanță precizarea limitelor respectivului interval. În cazul variabilelor discrete, prezente în exemplul nostru, limitele intervalului 32-37 sunt 32 și 37, ambele valori fiind incluse în această clasă (amintim că se obțin ca date experimentale doar valori întregi).

În cazul variabilelor continue care produc numere fracționare, este uzual să se indice limitele exacte, în cazul nostru 31,5-37,5, intervalele fiind închise la stânga și deschise la dreapta, [31,5; 37,5), cu excepția ultimului interval de clasă care este închis la ambele extremități. Și în acest caz se preferă uneori să se treacă în tabel limitele întregi, numite limite de raportare, dar în calcule se vor utiliza limitele exacte.

A treia rubrică prezintă mijlocul intervalului, obținut în cazul de față ca semisumă a limitelor intervalului, de exemplu 34,5 = (32+37)/2. S-a precizat anterior că această valoare constituie cea mai bună aproximare pentru orice dată experimentală care aparține intervalului. Cu alte cuvinte, se poate considera că orice persoană testată și care se încadrează în clasa 2 a obținut un scor de 34,5. Precizia aproximării crește dacă amplitudinea intervalelor este mai mică.

A patra coloană prezintă frecvențele claselor, deci numărul observațiilor încadrate în fiecare clasă. Poartă denumirea de frecvențe absolute și se notează cu f, utilizând indici pentru a identifica clasa. Frecvența totală, deci numărul de observații, desemnată prin N, este suma frecvențelor de clasă:

$$N = f_1 + f_2 + \ldots + f_k = \sum_{i=1}^{k} f_i ,$$

unde $f_i, i = 1, 2, \ldots, k$ desemnează frecvențele claselor, iar k numărul de clase. În cazul nostru N = 64, k = 9.

Rubrica a cincea cuprinde frecvențele relative de clasă, notate $\tilde{f}$. Frecvența relativă este o frecvență de clasă exprimată ca proporție a frecvenței totale: $\tilde{f}_i = f_i / N$. Suma tuturor frecvențelor relative este egală cu 1, fiecare fiind un număr cuprins între 0 și 1.

Este sugestivă exprimarea frecvențelor relative ca procentaje $\tilde{f}_i = (f_i/N)\times 100$, interpretarea fiind astfel facilitată. De exemplu, în tabelul 6.5, $f_2 = 7{,}8\%$ arată că, dintr-o sută de observații, aproximativ 8 se găsesc în clasa a doua. Este evident că suma frecvențelor relative exprimate procentual este egală cu 100.

Rubrica a șasea cuprinde frecvențele cumulate ascendent care se obțin prin însumarea progresivă a frecvențelor absolute:

$$Fc_i \downarrow = f_1 + f_2 + \ldots + f_i .$$

De exemplu, în tabelul 6.5,

$$Fc_5 \downarrow = f_1 + f_2 + \ldots + f_5 = 1+5+7+12+16 = 41 .$$

Interpretarea este imediată: există 41 de observații situate în primele 5 intervale (sau există 41 de observații care sunt mai mici sau egale cu 55, aceasta din urmă fiind limita superioară a intervalului). Frecvențele cumulate ascendent pot da răspuns la întrebări de genul „câte observații sunt mai mici decât...".

Frecvențele absolute se pot cumula și în ordinea inversă a enumerării intervalelor, obținându-se frecvențele cumulate descendent:

$$Fc_i \uparrow = f_k + f_{k-1} + \ldots + f_i ,$$

unde k desemnează numărul de intervale.

Deoarece are loc relația:

$$Fc_i \downarrow + Fc_{i+1} \uparrow = Fc_{i-1} \downarrow + Fc_i \uparrow = N ,$$

șirul frecvențelor cumulate descendent poate fi obținut din șirul frecvențelor cumulate ascendent prin:

$$Fc_i \uparrow = N - Fc_{i-1} \downarrow ,$$

adică din numărul total de observații se scade frecvența cumulată ascendentă corespunzătoare clasei precedente.

Coloana a șaptea din tabelul 4.7 conține frecvențele cumulate ascendent pentru exemplul prelucrat. Astfel,

$$Fc_5 \uparrow = f_9 + f_8 + \ldots + f_5 = 3+4+7+9+16 = 39 = N - Fc_4 \downarrow = 64-25.$$

Interpretarea este aceea că 39 de observații sunt mai mari sau egale cu 50 (limita inferioară a clasei 5). Deci frecvențele cumulate descendente răspund la întrebări de genul „câte observații sunt mai mari decât...". Frecvențele cumulate descendent mai sunt denumite *frecvențe retrocumulate*.

Rubricile 8 și 9 cuprind frecvențele cumulate relative ascendente și descendente, adică exprimarea rubricilor 6 și 7 sub formă de cote procentuale. Interpretările sunt imediate și reamintim că exprimările sub formă de proporții (procente) adimensionalizează mărimile transformate, ceea ce permite comparații între eșantioane sau studii diferite.

Observație. Numărul de rubrici atașate unei repartiții de frecvențe este în funcție de scopurile propuse. Astfel, dacă ne propunem să calculăm doar media, nu avem nevoie de frecvențele relative și cumulate.

În finalul capitolului prezentăm un indicator statistic mai puțin utilizat, și anume **frecvența unitară** sau **densitatea de frecvență absolută** a unui interval de grupare, definită prin raportul

$$\frac{f_i}{\Delta x_i},$$

unde Δx_i este mărimea intervalului *i* avut în vedere. Acest raport este utilizat mai ales în cazul alegerii unor intervale de grupare inegale şi are interpretarea imediată a numărului de observaţii care se distribuie pe o unitate de măsură, diferenţiind intervale mai „dense" sau mai „rare". Un indice analog poate fi calculat prin

$$\frac{\tilde{f}_i}{\Delta x_i}$$

pentru frecvenţele relative, cu o interpretare asemănătoare.

Deoarece exemplul 4.1 a fost prelucrat prin grupare în intervale egale, frecvenţele unitare nu aduc nici o informaţie în plus: este evident că, în situaţia grupării în clase egale, intervalele cu frecvenţe absolute (relative) mai mari au şi o frecvenţă unitară mai mare.

4.2. Protocoalele de bază ale distribuţiilor

Planul de conducere a unei experienţe şi a unei culegeri de date constituie un protocol. Prezentarea datelor conform unor protocoale este justificată de specificul cercetării psihologice, în care datele aparţin unor subiecţi supuşi unei forme de investigaţie. În primul caz, se vorbeşte de protocol experimental, care defineşte ansamblul condiţiilor experimentale, ordonarea lor temporală, influenţa condiţiilor experimentale asupra subiecţilor. Deoarece planul de experienţă reprezintă şi planul de culegere a datelor, noţiunea de protocol se aplică, de asemenea, observaţiilor care sunt culese.

Aşadar, protocolul datelor este un ansamblu de date structurate după un anumit plan, care constituie planul de experienţă. Acest plan condiţionează nivelul analizei datelor.

4.2.1. Protocolul de structură „subiect – rezultat total" (numit de Annick Weil Barais şi Edouard Frienel protocol de tip S_n)[30]

Fiecărui subiect s_i al unui ansamblu de n subiecţi, plasaţi într-o aceeaşi situaţie, îi corespunde o modalitate sau o valoare (notată x_i) a variabilei măsurate.

Exemplul 4.2. Rezultatul obţinut de elevii unei clase la testul *Matricea Progresivă Raven* (rezultat total). Exemplul este fictiv, rezultatele sunt ilustrate, pentru 106 subiecţi, în tabelul 4.8, secţiunea din dreapta. În partea stângă este dat tabelul teoretic. S reprezintă subiecţii, s_i subiectul i, x_i fiind scorul obţinut de subiectul respectiv; i variază de la 1 la n.

30. Annick Weil Barais (coord.), *Les méthodes en psychologie*, Collection Grand Amphi, Breal, Rosny, 1997, p. 188.

Tabelul 4.8. *Protocol de structură „subiect – rezultat total"*

S	x_i	S	x_i
s_1	x_1	1	23
s_2	x_2	2	34
s_3	x_3	3	27
–	–	–	–
s_n	x_n	106	45

Prelucrând tabelul 4.8, prin calcularea numărului n_i de observaţii (aici numărul de subiecţi) având o anumită valoare x_i, se obţine tabelul distribuţiei sau al repartiţiei de efective (frecvenţe absolute) în funcţie de modalităţile variabilei. Protocolul cuprinde, în acest caz, o distribuţie pe variante de variaţie. Fiecare valoare a lui x este diferită de alta şi suma efectivelor este egală cu numărul n (totalul observaţiilor), numit şi talia sau mărimea eşantionului. Rezultatul prelucrării este dat în tabelul 4.9.

Tabelul 4.9. *Distribuţia efectivelor corespunzând unui protocol de tip „subiect – rezultat total"*

x_i	Efective	x_i	Efective
x_a	n_a	23	4
x_b	n_b	27	12
x_c	n_c	32	23
–	–	–	–
x_k	n_k	67	6

În partea din dreapta a tabelului 4.9 se observă că 4 subiecţi au obţinut rezultatul 23, 12 subiecţi au rezultatul 27 etc.

Dacă se divizează frecvenţele absolute prin efectivul total, n, se obţine un tabel al distribuţiei frecvenţelor relative, în care suma acestora are valoarea 1 (valoarea unei frecvenţe relative este cuprinsă între 0 şi 1). Se obţine astfel tabelul 4.10:

Tabelul 4.10. *Distribuţia de frecvenţe relative corespunzând unui protocol de tip „subiect – rezultat total"*

x_i	Frecvenţa	x_i	Frecvenţa
x_a	f_a	23	0,037
x_b	f_b	27	0,110
x_c	f_c	32	0,210
–	–	–	–
x_k	f_k	67	0,057

Reamintim că frecvenţele relative pot fi exprimate procentual.

Punctajul obţinut de un subiect la un test psihologic se numeşte **scor**. Acelaşi nume îl poartă şi rezultatul unui item, adică la un „punct" din test. Rezultatul total al unui test mai poartă denumirea de **scor compozit**, fiind obţinut prin însumarea

scorurilor itemilor. Rezultatele unor categorii de teste nu pot fi separate în itemi. Este cazul testului de asociere şi disociere manuală conceput de J.M. Lahy, care nu poate fi descompus în itemi. Nu este singurul exemplu.

Între scorurile de la itemi şi scorurile compozite există o relaţie foarte strânsă. Are importanţă, mai ales, modul cum se raportează varianţa scorului compozit la varianţele itemilor. Etalarea unui scor pune probleme de precizie matematică şi probleme de conţinut. Ultimele se referă la ce măsurăm de fapt prin adunarea itemilor, deoarece aceştia pot să se refere la caracteristici diferite, mai ales când testele au la rândul lor subteste, cum este cazul la WISC şi la chestionarele de personalitate 16PF, MMPI. Aceste subteste sau scale (uneori, pentru denumirea testului în întregime se foloseşte termenul de „scală") au, la rândul lor, scoruri compozite. Problemele de conţinut fac obiectul preocupării predominante a disciplinei numite psihometrie.

În funcţie de profunzimea analizei dorite, o importanţă deosebită prezintă modul în care înregistrăm scorurile compozite, în ansamblu sau cu înregistrarea detaliată a rezultatelor itemilor. Pierderile de informaţii sunt foarte mari în cazul în care înregistrăm doar rezultatul total, iar nevoia de utilizare a unor calcule statistice de un rafinament superior nu va putea fi satisfăcută. Aceste calcule sunt particularizate de teoria modernă a răspunsurilor la itemi. Dacă avem posibilitatea să prezentăm detaliat modul de răspuns la un item, putem pune în aplicaţie o gamă diversificată de procedee statistice.

Exemplul 4.3. Un set de date rezultat în urma examinării a 12 subiecţi, prin intermediul unui test compus din 12 itemi (fiecare avea 8 modalităţi de răspuns la alegere, dintre care numai una era corectă), poate fi prezentat în următoarele moduri (exemplul imaginat este de la o secţiune a Matricei Progresive Raven). Într-o primă variantă putem prezenta rezultatele înregistrate global, luând în considerare şi răspunsurile greşite.

Tabelul 4.11. *Rezultate globale (exemplul 4.2)*

Subiectul	Răspunsuri corecte	Răspunsuri greşite
1	6	6
2	8	4
3	9	3
4	6	6
5	5	7
6	3	9
7	5	7
8	8	4
9	11	1
10	12	0
11	7	5
12	2	10

Asupra datelor oferite în acest fel putem efectua operaţii de calcul al indicatorilor statistici (indicatori de nivel, indicatori de împrăştiere, indicatori de asimetrie şi indicatori de exces) şi putem, în funcţie de respectivii indicatori, să stabilim etaloane condiţionate de particularităţile distribuţiei valorilor. Dar cu aceste date nu putem reevalua constructiv testul, în cazul în care ne propunem aşa ceva.

Un alt mod de prezentare îl constituie cel în care rezultatul este descompus pe itemi. Valoarea 1 indică rezolvarea unui item, iar valoarea 0 nerezolvarea acestuia. Se alcătuieşte astfel matricea item – subiect, stabilită de W. Stern în 1911 (cunoscută şi sub titulatura de matrice Stern). Numărul itemilor este trecut în prima linie, iar numărul subiecţilor în prima coloană. Liniile 1, 2, 3, 4 ... conţin rezultatul subiecţilor la itemi, iar coloanele 1, 2, 3, 4 ..., rezultatele obţinute la un item de către cei 12 subiecţi. În matricea subiecţi – itemi se vorbeşte de standardizare când calculele cuprind itemii şi de ipsatizare când calculele cuprind subiecţii.

Tabelul 4.12. *Matricea subiecţi – itemi (exemplul 4.2)*

Item / Subiect	1	2	3	4	5	6	7	8	9	10	11	12	C	G
1	1	1	1	1	0	0	1	1	0	0	0	0	6	6
2	1	1	1	1	0	1	1	1	0	0	1	0	8	4
3	1	1	1	1	1	1	1	0	1	0	1	0	9	3
4	1	1	1	1	1	1	0	0	0	0	0	0	6	6
5	1	1	0	0	0	0	0	0	1	1	1	0	5	7
6	0	0	0	0	0	0	1	1	1	0	0	0	3	9
7	1	1	1	0	1	1	0	0	0	0	0	0	5	7
8	1	1	1	1	0	0	0	1	1	1	0	1	8	4
9	1	0	1	1	1	1	1	1	1	1	1	1	11	1
10	1	1	1	1	1	1	1	1	1	1	1	1	12	0
11	1	1	1	0	1	1	1	1	0	0	0	0	7	5
12	1	0	1	0	0	0	0	0	0	0	0	0	2	10
C	11	9	10	7	6	7	7	7	6	4	5	3		
G	1	3	2	5	6	5	5	5	6	8	7	9		

Legendă: C = răspunsuri corecte, G = răspunsuri greşite.

În ultimele două coloane sunt înregistrate totalurile răspunsurilor corecte/greşite pentru fiecare subiect în parte. În ultimele două linii sunt înregistrate totalurile răspunsurilor corecte/greşite pentru fiecare item. Astfel, se poate observa că subiectul al treilea a dat nouă răspunsuri corecte şi trei greşite, în timp ce la primul item 11 subiecţi au răspuns corect şi doar unul a dat un răspuns greşit.

Dacă avem la dispoziţie un astfel de set de date, putem efectua operaţiunile statistice amintite anterior, plus o serie de operaţiuni specifice acestui mod de prezentare. Astfel, pentru un asemenea set de date se pot efectua următoarele operaţii:

- calculul indicelui de dificultate a itemului, putându-se indica în acest caz itemii care nu corespund, conform exigențelor teoriei clasice a testului;
- calculul indicelui de discriminare, necesar, de asemenea, eliminării din test a unor itemi;
- calculul matricei de corelație interitemi, calculul coeficienților de corelație item – rezultat final și calculul matricei de varianță – covarianță. Aceste calcule sunt strict necesare pentru stabilirea omogenității itemilor, pentru calibrarea acestora și, în cazul disponibilității datelor pentru un criteriu, pentru stabilirea itemilor supresori (itemi care sunt corelați cu rezultatul final al testului, dar nu se corelează cu criteriul respectiv).

Același set de date poate fi prezentat și sub următoarea formă:

Tabelul 4.13. *Matricea subiecți – itemi, răspunsuri (exemplul 4.2)*

Item / Subiect	1	2	3	4	5	6	7	8	9	10	11	12	C	G
1	8	2	3	8	1	3	5	1	1	5	3	6	6	6
2	8	2	3	7	7	4	5	1	6	5	1	3	8	4
3	8	2	3	8	7	4	5	2	7	5	1	5	9	3
4	8	2	3	8	7	4	4	5	6	9	3	5	6	6
5	8	2	4	5	6	3	2	4	7	6	1	3	5	7
6	7	3	5	4	4	2	1	5	1	7	4	5	3	9
7	8	2	3	6	7	4	2	8	5	3	4	5	5	7
8	8	2	3	6	6	2	3	1	7	6	1	5	8	4
9	8	4	3	8	7	4	5	1	7	6	1	4	11	1
10	8	2	3	8	8	7	4	5	7	6	1	4	12	0
11	8	2	3	7	5	3	7	8	7	6	1	4	7	5
12	8	3	3	7	5	2	8	2	3	5	4	3	2	10
C	11	9	10	7	6	7	7	7	6	4	5	3		
G	1	3	2	5	6	5	5	5	6	8	7	9		
VC	8	3	3	7	5	2	8	2	3	5	4	3		

Legendă: C = răspunsuri corecte, G = răspunsuri greșite, VC = varianta corectă.

Față de legenda anterioară, la acest tabel avem adăugat VC, desemnând varianta corectă de răspuns. Spre deosebire de tabelul 4.12, în care răspunsurile erau prezentate sub forma dihotomică (0 și 1), în acest caz casetele tabelului conțin varianta la care a răspuns subiectul, indiferent dacă este corectă sau incorectă. În acest fel putem observa proporția diferitelor variante de răspuns în raport cu numărul total de subiecți luați în calcul. Având rezultatele trecute în acest mod, putem efectua toate calculele amintite anterior și vom avea la dispoziție variantele de răspunsuri greșite. Considerăm că acest mod de analiză a rezultatelor testului poate purta numele de infraanaliză (cu sensul de analiză „în interiorul" testului, analiză „sub" test, produsă prin descompunerea testului). Nivelul grosier, global al rezultatului ar putea constitui un suport informațional pentru un nivel infraanalitic 0, datele prezentate cu etalarea rezultatului sub forma 0 și 1 ar putea constitui un

suport informațional pentru un nivel infraanalitic 1, iar ultima variantă pentru un nivel infraanalitic 2. Dacă am adăuga și calcularea timpului de răspuns pentru fiecare variantă, am constitui un nivel infraanalitic 3. Nivelul 1 poate fi identificat cu analiza de itemi. Informații suplimentare vom oferi într-un capitol ulterior, dedicat entropiei și distractorului.

4.2.2. Protocolul de structură „subiect – grup de apartenență" (denumit de autorii amintiți anterior protocol de structură S<G>)[31]

Datele obținute provin de la subiecți aparținând unor grupuri diferite sau care au fost plasați în situații diferite și constituie serii independente de măsuri, denumite și grupuri sau eșantioane independente.

Tabelul 4.14. *Protocol de bază de tipul „subiect – grup de apartenență"*

S	G	x_i
s_1	g_1	x_1
s_2	g_1	x_2
–	–	–
s_i	g_i	x_j
–	–	–
–	–	–
s_n	g_q	x_n

S	G	x_i
1	grup 1	25
2	grup 1	34
–	–	–
23	grup 2	45
–	–	–
–	–	–
118	grup 7	65

Legenda: S = subiect, G = grup, X_i = rezultat, n = numărul de subiecți, q = numărul de grupuri.

În acest caz se construiește o distribuție prin grup care indică numărul de observații corespunzând diferitelor valori ale variabilei măsurate (în fiecare dintre grupuri).

Dacă numărul de observații nu este același în fiecare grup, altfel spus, dacă gruparea nu este echilibrată, se impune prezentarea distribuției frecvențelor relative. Următoarele trei tabele conțin situația generală, exemplificarea frecvențelor absolute și, respectiv, exemplificarea frecvențelor relative.

Tabelul 4.15. *Situația generală a frecvențelor absolute*

Grupuri	Valori ale variabilei măsurate				
	x_a	x_b	–	x_k	Total
g_1	n_{a1}	n_{b1}	–	n_{k1}	n_1
g_2	n_{a2}	n_{b2}	–	n_{k1}	n_2
–	–	–	–	–	–
g_q	n_{aq}	n_{bq}	–	n_{kq}	n_q
Total	n_a	n_b	–	n_k	n

31. Annick Weil Barais, *op.cit.*, p. 188.

Tabelul 4.16. *Distribuția frecvențelor absolute în funcție de variantele de răspuns și de grupuri*

Grupuri	Valori ale variabilei măsurate				
	x_a	x_b	–	x_k	Total
g_1	5	7	–	15	34
g_2	16	5	–	15	42
–	–	–	–	–	–
g_q	10	4	–	17	37
Total	32	25	–	35	118

Fiecare linie din tabelul 4.16 conține distribuția scorurilor unui grup după variantele de răspuns. Tabelul efectivelor se completează cu efective „de margine" sau „marginale", denumite și totaluri parțiale. Fiecare coloană a tabelului conține distribuția după grupuri a subiecților care au obținut un același scor (notat în antetul coloanei). Ultima coloană conține distribuția subiecților după grupuri (indiferent de scorul obținut), iar ultima linie – distribuția subiecților după scorurile obținute (indiferent de grupul din care fac parte). Din tabelul 4.16 se poate citi, de exemplu, că în grupul 1 există 5 subiecți care au obținut scorul x_a, iar în grupul q există 17 subiecți care au obținut scorul x_k.

Tabelul 4.17. *Distribuția frecvențelor relative în funcție de variantele de răspuns și de grupuri*

Grupuri	Valori ale variabilei măsurate				
	x_a	x_b	–	x_k	Total
g_1	0,14	0,20	–	0,35	1,00
g_2	0,38	0,11	–	0,35	1,00
–	–	–	–	–	–
g_q	0,27	0,10	–	0,45	1,00
Total	0,27	0,21	–	0,45	1,00

Interpretând tabelul 4.17, în grupul 1 o proporție de 0,14 dintre subiecți au scorul x_a, iar în grupul q o proporție de 0,45 dintre subiecți au scorul x_k.

Frecvențele relative sunt calculate prin divizarea efectivului obținut în fiecare grup, pentru fiecare variantă a variabilei, la mărimea grupului respectiv. Operația realizată se numește calculul frecvențelor relative „în linie". Se ține cont și de mărimea respectivă a eșantioanelor, dar și de numărul total al observațiilor pentru fiecare variantă a variabilei măsurate.

Atragem atenția asupra faptului că linia de total nu se obține prin însumarea frecvențelor relative de pe coloanele respective, ci este obținută prin raportarea valorilor din linia „total" a tabelului 4.16 la numărul total de subiecți.

Pentru acest format de tabel există trei tipuri diferite de frecvențe relative:

- „în linie" sau „pe linie", când se divizează fiecare frecvență absolută prin totalul liniei respective (suma unei linii, a unui grup, produce valoarea 1,00). Însumarea pe coloană nu are sens. O exemplificare în acest sens prezintă subtabelul al doilea din tabelul 5.
- „în coloană" sau „pe coloană", când se divizează fiecare frecvență absolută prin totalul coloanei respective (suma unei coloane, a unei modalități a variabilei măsurate, are valoarea 1,00). Însumarea pe linie nu are sens.
- „încrucișate", când se divizează fiecare frecvență absolută prin totalul general n. Suma tuturor frecvențelor relative interne are valoarea 1,00, ca și suma frecvențelor marginale, atât în linie, cât și pe coloană.

Toate cele trei modalități de operare sunt pertinente, oferind viziuni diferite asupra structurii rezultatelor.

4.2.3. Protocolul de structură „subiect – condiție de investigație"

(denumit de Annick Weil Barais și Edouard Frienel „protocol de structură S*T")[32]

Acest fel de protocoale se realizează în cazul când subiecții sunt supuși la mai multe condiții (T). Există repetări ale măsurătorilor și seriile de măsurători efectuate în fiecare modalitate a factorului T sunt perechi, suportate de aceiași subiecți.

Tabelul 4.18. *Protocol de bază de tipul „subiect – condiție de investigație"*

S \ T	t_1	t_2
s_1	x_{11}	x_{12}
s_2	x_{21}	x_{22}
...	...	...
s_p	x_{p1}	x_{p2}

S \ T	t_1	t_2
1	34	45
2	36	42
...	...	...
170	21	47

Legendă: S = subiecți; t_1, t_2 = modalități ale factorului T.

De exemplu, subiecții răspund la mai multe întrebări, sunt confruntați cu diferiți stimuli și sunt plasați în diferite condiții. Situația poate fi ilustrată printr-un studiu longitudinal asupra capacității de concentrare a atenției la începutul și la terminarea unei zile de lucru, în care a fost efectuată aceeași probă la începutul și la încheierea timpului de lucru. (T are în acest caz două modalități.) Aceste date experimentale se pot examina în maniere diferite:

- se compară performanțele la începutul și la sfârșitul timpului de lucru;
- se analizează distribuțiile pentru fiecare modalitate a factorului T;
- se studiază relațiile între probe.

Reprezentarea cea mai adecvată a distribuției de efective (frecvențe relative sau procentaje) o constituie aceea care ține cont de relația de încrucișare dintre S și T.

32. Annich Weil Barais (coord.), *op.cit.*, p. 189.

Tabelul 4.19. *Situația generică a organizării datelor în protocolul de tip „subiect – condiție de investigație"*

T	t_2					Total
	X	x_a	x_b	...	x_k	
t_1	x_a	n_{aa}	n_{ab}	...	n_{ak}	n_{a1}
	x_b	n_{ba}	n_{bb}	...	n_{bk}	n_{b1}
	...	...	...	...	...	...
	x_k	n_{ka}	n_{kb}	...	n_{kk}	n_{k1}
Total		n_{a2}	n_{b2}	...	n_{k2}	n

Citirea/organizarea unui asemenea tabel se face după modelul: n_{aa} desemnează numărul de subiecți care au obținut același scor x_a în ambele modalități ale factorului T (notate t_1 și t_2). Fiind implicate frecvențele absolute, are sens calcularea sumelor de frecvențe atât pe linie, cât și pe coloană, obținându-se astfel distribuțiile marginale. Astfel, n_{a1} reprezintă numărul de subiecți care, în modalitatea t_1, au obținut scorul x_a. În mod analog, n_{b2} reprezintă numărul de subiecți care în modalitatea t_2 au obținut scorul x_b. Distribuțiile marginale constituie repartițiile scorurilor în fiecare modalitate. Aceste două distribuții parțiale permit compararea între modalitățile t_1 și t_2.

O exemplificare a tabelului generic 4.19 este dată în tabelul 4.20:

Tabelul 4.20. *Exemplu de organizare a datelor în protocolul de tip „subiect – condiție de investigație"*

T	t_2					Total
	X	16	19	...	34	
t_1	16	12	17	...	20	97
	19	15	19	...	11	87
	...	...	...	...	...	...
	34	19	22	...	24	104
Total		110	112	...	216	540

Conform explicațiilor din situația generală, putem ști că 12 subiecți au obținut scorul 16 în ambele modalități T, 97 de subiecți au obținut scorul 16 în modalitatea t_1, iar 112 subiecți au obținut scorul 19 în modalitatea t_2.

Calculând frecvențele relative prin raportarea la totalul general, se obține tabelul 4.20, care arată ponderea relativă a fiecărei casete din distribuția comună a celor două modalități T.

Tabelul 4.21. *Distribuția frecvențelor relative prin raportare la numărul total de subiecți*

T	t_2					Total
	X	16	19	...	34	
t_1	16	0,02	0,03	...	0,04	0,18
	19	0,027	0,035	...	0,02	0,16
	...	...	...	...	...	...
	34	0,035	0,04	...	0,04	0,19
Total		0,20	0,20	...	0,40	1,00

Interpretarea valorilor din tabel se face după modelul: 0,02 reprezintă proporția subiecților (din totalul de 540 de subiecți) care au scorul 16 în cele două modalități T, 0,20 reprezintă proporția subiecților care au obținut scorul 19 în t_2, iar 0,18 proporția subiecților care au obținut scorul 16 în t_1.

Tabele similare tabelului 4.21 se pot forma prin raportarea frecvențelor absolute din casetele interioare la totalurile pe linie sau pe coloane. Interpretarea unor asemenea prelucrări se face după modelul amintit la protocolul „subiect – grup de apartenență", tabelul 4.17.

Capitolul 5

NOȚIUNI DE TEORIA PROBABILITĂȚILOR

Capitolele anterioare arată că interesul unui cercetător este acela de a cunoaşte caracteristicile unei populaţii pe care am denumit-o populaţie statistică. Termenul trebuie înţeles în sensul precizat atunci (de exemplu, o populaţie statistică desemnează mulţimea reacţiilor posibile ale unui individ la un anumit stimul sau elevii din clasa a V-a ori locuitorii unui oraş etc.). Dar, din diverse motive, sunt foarte rare cazurile când un cercetător poate studia întreaga populaţie şi poate recurge la eşantioane reprezentând populaţia vizată. Rezultatele obţinute din analiza eşantionului sunt generalizate pentru întreaga populaţie.

Atunci când se efectuează o asemenea generalizare, apar în mod natural două întrebări: dacă operaţia produce un rezultat credibil şi câtă încredere putem avea în rezultatele extinse asupra întregii populaţii. Pentru a răspunde acestor întrebări, este nevoie de noţiunile prezentate în capitolul de faţă. Apar noţiunea de probabilitate şi noţiunile conexe necesare formării unui mod de raţionare corect.

Trebuie să se observe faptul că termenii *credibil* şi *încredere*, utilizaţi în paragraful precedent, au un înţeles încărcat de subiectivism şi, prin urmare, este necesară o formalizare a lor.

5.1. Modele teoretice

5.1.1. Raportul teorie – model

Este necesar să facem precizări asupra acestui raport, deoarece ignorarea lor şi izolarea statisticii de contextul şi specificul cercetării duc la o serie de acuzaţii nedrepte în privinţa acestei discipline. Intervenţia procedurilor statistice nu este un act ştiinţific specializat izolat, ci face parte dintr-un şir de etape ale demersului ştiinţific. Este adevărat că procedurile statistice nu sunt utilizate doar de cercetători specializaţi, ci şi de psihologi practicieni care au preocupări proprii vizând relevanţa datelor.

În multe discipline ştiinţifice din secolul XX se acordă o foarte mare importanţă teoriilor, în aşa fel încât s-au creat deja „teorii ale teoriilor" sau metateorii. În sens ştiinţific, teoria integrează mai multe relaţii între fapte care până la acest demers erau considerate independente. Pe lângă o astfel de contribuţie la înţelegerea ansamblului de fapte, teoria generează noi ipoteze de cercetare şi îndeplineşte o funcţie euristică esenţială. Cunoscutul filosof Mario Bunge compară teoriile cu

prezenţa unui sistem nervos, mai mult sau mai puţin elaborat, la fiinţele vii. Teoriile constituie deci sistemul nervos al unei ştiinţe. Cunoştinţele ştiinţifice sunt fapte „preparate" cu ajutorul teoriilor. Faptele rămân inutile şi nefolositoare atâta vreme cât nici o teorie nu le adună şi nu le organizează într-un tot coerent. Teoriile permit prezicerea, controlarea şi explicarea, în unele cazuri.

Psihologia este o disciplină caracterizată printr-o extensie explozivă a cantităţii de date produse şi care realizează generalizări empirice. Însă această disciplină este considerată încă, din punct de vedere teoretic, o disciplină „embrionară". „Culegerea de date şi descrierea rămân în acest caz activităţile cele mai respectabile, într-o asemenea măsură încât, în acest domeniu, mai mulţi cercetători opun adesea acumularea de date elaborării teoriilor, pe care le consideră simple speculaţii. Această atitudine ignoră faptul că datele nu capătă o semnificaţie decât într-un context teoretic şi că acumularea haotică de fapte – chiar acumularea sistematică de generalizări empirice – este o pierdere de timp dacă nu este ghidată prin teorie"[33], afirmă Michel Robert, profesor la Universitatea din Montréal, în spiritul concepţiei lui Mario Bunge.

Teoriile sunt produsul activităţii ştiinţifice şi punctul lor de plecare este constituit din enunţuri şi ipoteze; o asemenea afirmaţie este, cel puţin parţial, falsă, deoarece astfel are loc o simplificare a realului. Câtă vreme se află în „vogă", o teorie reprezintă un instrument ştiinţific dinamic care ghidează observaţia, se modifică pentru a se adapta faptelor, tinde să reducă incoerenţa; în final însă este condamnată la desuetudine, din momentul în care apare o altă teorie, mai interesantă şi mai înglobantă.

Din punct de vedere metodologic, calitatea esenţială a unei teorii ştiinţifice este aceea că poate fi, direct sau indirect, confirmată de datele empirice. Ea trebuie să fie în măsură să genereze predicţii sau ipoteze, putând constitui probă, mai exact coroborată sau infirmată prin experienţă. Nu poate fi considerată ştiinţifică teoria care nu poate face proba respingerii sale (Karl Popper).

5.1.2. Modele în psihologie

Pentru obiectul de interes al acestei lucrări, este necesară o scurtă introducere în lumea modelelor, deoarece în psihologie, şi nu numai aici, măsurarea are sens doar prin raportare la un model.

Cuvântul „model" provine din cuvântul latin *modus*, celelalte limbi preluându-l prin intermediul cuvântului italian *modello* (diminutiv al lui *modus*). La început, termenul a fost folosit în arta plastică, ulterior în tehnică şi apoi în diverse discipline ştiinţifice aplicative. Folosirea lingvistică acoperă în mare parte semnificaţia iniţială a cuvântului: „mod sau fel diminuat (micşorat, redus) al unui obiect" (al unei realităţi).

Folosirea termenului „model" are sensuri diferenţiate în domeniul psihologic.

33. Michel Robert (coord.), *Fondements et étapes de la recherche scientifique en psychologie*, Edisem, Maloine S.A., Paris, 1988, p. 19.

În sens general statistic, modelul este „expresia sub o formă logică şi matematică a proprietăţilor pe care trebuie să le prezinte un anumit tip de date, dacă aceste date se conformează exact la o anumită ipoteză sau teorie pe care modelul o formalizează"[34], de exemplu modelul unei scale ordinale sau cel al unei variabile normale.

În sens strict statistic, modelul este „prezentarea, sub forma unei ecuaţii sau a unui sistem de ecuaţii, a relaţiilor teoretice aşteptate între mai multe variabile aşteptate şi latente"[35]. Unele dintre ele sunt variabile ce urmează să fie explicate (variabile dependente, criterii), altele sunt variabile explicative (variabile independente, predictori), în timp ce altele sunt variabile aleatorii (erori) reprezentând efectul eventual al variabilelor care nu figurează în model. Exemple în acest sens, mai restrâns, constituie ecuaţiile de regresie simple şi multiple, ecuaţiile reprezentând o variabilă observată în funcţie de variabile latente, ca în analiza factorială, şi ecuaţiile unei analize în piste cauzale.

În psihologia dezvoltării, modelul este o referinţă comportamentală a copilului sau a adultului. Modelul desemnează în acest caz pe cel care este imitat.

Psihologia ca domeniu de investigaţie ştiinţifică are ca primă funcţie evidenţierea fenomenelor empirice, în condiţii de control şi de reproducţie optimală. A doua funcţie a psihologiei ştiinţifice este cea de a propune explicaţii, permiţând coordonarea faptelor şi prezentarea în această privinţă a unei viziuni sintetice, economice şi euristice. Construcţia modelelor psihologice tinde să realizeze această funcţie.

Printr-un model se încearcă să se prezinte esenţa unei realităţi într-un mod mai lapidar şi mai simplificat. Ceea ce, în acest sens, este „esenţa" unei realităţi depinde, natural, de scopul construcţiei modelului. Exemplu: într-un anumit fel este prezentat modelul unui automobil dacă scopul este acela de a-l face atractiv pentru cumpărători şi în cu totul alt fel dacă scopul este de a-l prezenta unui cerc foarte restrâns de specialişti, preocupaţi doar de rezistenţa la tamponare a respectivului vehicul.

Un model este un discurs logico-matematic asupra realităţii, care nu poate fi prezentat tot timpul prin intermediul unei scheme. Desigur, se poate reprezenta caracterul esenţial al unui model sub formă schematică şi numeroase tipuri de discursuri pot fi schematizate, chiar şi cele aparţinând metafizicii şi fenomenologiei. De asemenea, conceptul de model nu poate fi echivalat cu cel de teorie, aceasta din urmă constituind un câmp ideatic mai vast, mai abstract. Există nenumărate teorii ştiinţifice care nu pot fi reduse la modele ştiinţifice, dar care sunt de aceeaşi natură şi întreţin cu ele relaţii evidente. Ceea ce diferenţiază teoriile de modele este în primul rând gradul lor de generalitate: un model este o reprezentare logico-matematică a unui ansamblu limitat de fenomene în condiţii atent definite; o teorie ştiinţifică se aplică însă la un ansamblu mult mai vast de fenomene. O teorie ştiinţifică poate fi considerată şi un model de ordin superior, din care pot fi derivate

34. Vezi *Grand dictionnaire de la psychologie*, Larousse, Paris, 1992, p. 474.
35. *Ibidem*.

modele locale, adică realizări specifice care se aplică la anumite obiecte particulare şi la anumite clase definite de situaţii. Mecanica clasică vizează o teorie din care se pot deriva modele de explicaţie locală a mişcării pendulului, a dinamicii mareelor sau a mişcării planetelor. Integrarea modelelor locale într-o teorie generală constrânge la a considera unele fenomene diferite ca fiind fundamental identice, ceea ce impune, desigur, introducerea unui număr mare de postulate. Elaborarea unei teorii ştiinţifice este însoţită, aşadar, de o deplasare spre grade din ce în ce mai mari de abstractizare, oferind din ce în ce mai multă experimentare critică şi decisivă. Cu cât teoria selecţionează în mod necesar un număr mai mare de modele locale, care pot fi integrate, cu atât mai multe dezbateri teoretice rezultă frecvent din incompatibilitatea modelelor de bază luate în considerare. Un model local selecţionează mult, dar generalizează puţin, în timp ce o teorie selecţionează puţin, dar generalizează mult.

În psihologie, numărul de modele este considerabil, dar, în schimb, nu există decât puţine teorii care să fie realmente şi total satisfăcătoare dintr-un punct de vedere strict logico-matematic. Modelul nu este o teorie, ci o microteorie, ceea ce-i dă câteva avantaje din punctul de vedere al preciziei, dar care limitează din această pricină capacităţile de generalizare.

Modelul formal rezultă din aplicarea unui sistem formal la un subansamblu de obiecte ale realităţii, în scopul de a permite o interpretare în termeni de valoare de adevăr. Ca atare, un model trebuie să posede o sintaxă (aceea a sistemului formal utilizat) şi o semantică (rezultând din proiecţia unui sistem formal într-o realitate semnificantă). Sintaxa sistemului formal este definită printr-un ansamblu de condiţii (un limbaj) şi presupune definirea simbolurilor elementare ale sistemului, ale expresiilor sistemului, ale regulilor de formare a expresiilor şi ale expresiilor corect formulate ale sistemului (formule din sistem). Un sistem formal trebuie, de asemenea, să definească o procedură de demonstraţie formalizată (axiome, relaţii între formule şi reguli de inferenţă, teoreme). Sintaxa unui sistem formal include, aşadar, formalizarea limbajului şi formalizarea demonstraţiei. Elaborarea unui sistem formal este deci o operaţie de tip logic şi matematic. Pentru psihologie, problema esenţială este aceea a aplicării unui sistem formal la un obiect particular (inteligenţă, memorie etc.). Semantica trimite la interpretarea elementelor sistemului formal, adică la un model, indiferent dacă este psihologic sau nu; acesta posedă o sintaxă, dar şi o semantică, adică o interpretare. Construcţia unui model presupune un decupaj particular al realităţii studiate, care nu atinge regulile formale, şi o economie a ansamblului programului ştiinţific (incluzând determinante istorice, sociologice, precum şi înclinaţii proprii cercetătorului). Se poate întâmpla ca sintaxa unui sistem formal să fie aplicată într-un mod parţial sau chiar în întregime eronat într-un model psihologic, dar această dificultate poate fi, în principiu, evitată. Interesul unui sistem formal este că el utilizează o sintaxă, de ordin logic sau matematic, construită pentru descrierea şi interpretarea unui sector al realităţii. Guy Tiberghien face o precizare foarte importantă în legătură cu necesitatea formalizării: „Dacă interfaţa între o sintaxă a unui sistem formal şi

semantica unui model este o sursă de probleme considerabile, o analiză [agramaticală] a realului va avea consecinţe încă şi mai grave. Cu cât modelul are mai multe funcţii pozitive, cu atât va fi mai uşor de subestimat. În ciuda caracterului său semantic aproximativ, un model formal are avantajul de a fi specific (în privinţa unei situaţii), explicit (privind unele postulate şi axiome), condensator (în raport cu complexitatea datelor) şi, în sfîrşit, modificabil, prin schimbarea postulatelor sale"[36].

Un model formal poate fi definit ca un sistem relaţional, de formă, de structură şi de funcţie între două obiecte, dintre care unul este un sistem formal, în timp ce altul priveşte ansamblul limitat de obiecte pe care cineva doreşte să-l descrie şi/sau să-l studieze. În mod frecvent, sunt comparate două obiecte naturale între ele şi se sugerează că unul este modelul celuilalt (modelul animal al învăţării umane). Este, de asemenea, posibil să se compare un obiect artificial cu un obiect natural (modelul informatic al memoriei umane). Cele două tipuri de comparaţii nu sunt echivalente, deoarece în primul caz nu se dispune de sistemul formal de referinţă, pe când obiectul artificial al celei de-a doua comparaţii este realizarea materială a unui sistem formal (ordinatorul de tip von Neumann reprezintă o realizare materială a maşinii formale a lui Turing). Astfel de diferenţe nu sunt cu certitudine anodine, pentru că ele determină în mare parte dezbaterile actuale privind validitatea modelului informatic în psihologie, a proprietăţilor de „calculabilitate" ale sistemului cognitiv uman şi, în fine, privind posibilitatea „emulaţiei" prin sisteme artificiale de tratament al informaţiei.

Există, într-o bună parte a lumii psihologilor, o aversiune nedisimulată faţă de exprimarea formalizată şi mai ales faţă de cea cantitativă în psihologie. Acest fel de abordare a faptului psihologic ar reduce mult, într-o asemenea optică, din nuanţarea cercetării şi ar „tehniciza" excesiv exprimarea în această disciplină umanistă.

Autorul german Harald Meyer remarcă existenţa unei astfel de aversiuni şi consacră chiar un studiu acestui aspect. Cauzele respectivei stări de fapt se află în:

- aversiunea faţă de matematică şi statistică;
- aversiunea faţă de unele demersuri concrete cu metode matematico-statistice;
- neîncredere în sensul şi utilitatea metodelor matematico-statistice în domeniile de specialitate corespunzătoare[37].

În partea opusă se situează cei care le reproşează psihologilor „literari", caracterul vag şi imprecis al exprimării constatărilor din domeniul investigaţiei ştiinţifice, inadecvarea la cursul actual al ştiinţelor psihologice în contextul

36. Guy Tiberghien, „Modèles de l'activité cognitive", în Jean-Paul Caverni; Claude Bastien; Patrick Mendelsohn; Guy Tiberghien: *Psychologie cognitive. Modèles et méthodes*, Presses Universitaires de Grenoble, Grenoble, 1991.
37. Harald Von Meyer, „Statistikangst und Einstellungen gegenüber Statistik bei Studierenden der Verhaltenwissenschaften", în *Forschungsforum, Berichte aus Otto-Friedrich-Universität Bamberg „Mit Leib und Seele" Psychologie in Bamberg*, editată de Von Lothar Laux şi Hans Reinecker, Bamberg, 1994, pp. 67-73.

familiilor ştiinţifice din care face parte psihologia, care impune schimbarea inventarului conceptual şi tehnicizarea acestuia. Psihologia cognitivă, în mare progres metodologic, dă chiar un avertiment în această privinţă. Mircea Miclea susţine că psihologia resimte două feluri de presiuni: unele de sus în jos şi altele de jos în sus. În concepţia profesorului universitar clujean, presiunea de sus ar veni din partea specialiştilor în domeniul inteligenţei artificiale, care, în tentativa lor de construire a sistemelor inteligente, şi-au exprimat clar totala insatisfacţie faţă de oferta ştiinţifică a psihologilor tradiţionalişti. Conceptele clasice, facultaţioniste, sunt considerate de unii reprezentanţi ai acestui curent „comice şi naive" (Naraznan, 1986), „incapabile să ofere soluţii practice pentru realizarea sistemelor de inteligenţă artificială"[38]. Presiunea de jos provine din partea neuroştiinţelor, discipline care au cunoscut în ultimele decenii o intensă dezvoltare. „Descoperirea neurotransmiţătorilor, a neuromodulatorilor, a neurohormonilor, a sinapselor electrice, utilizarea tomografiei computerizate, magnetoencefalografia etc. sunt numai câteva din roadele acestei revoluţii, aflate în plin proces de desfăşurare. Ca urmare a acestor descoperiri s-a creat un decalaj enorm între nivelul de analiză infracelular, extrem de detaliat, practicat de neuroştiinţe şi analiza molară în termeni facultaţionişti, practicată de psihologi. Frustraţi de lipsa de corespondenţă dintre datele neurobiologice şi cele psihologice, specialiştii din neuroştiinţe au somat psihologii să-şi rafineze aparatul conceptual şi nivelul de analiză a proceselor psihice, altfel vor fi nevoiţi să reinventeze psihologia (Bunge, Ardila, 1987)"[39].

Un punct de vedere înţelept în privinţa modurilor de comunicare a rezultatelor cercetării este exprimat de Vasile Pavelcu într-o lucrare clasică a psihologiei româneşti, în care întâlnim o intersecţie fericită între cultura de specialitate, rafinamentul expresiv şi profunzimea analitică şi sintetică: „Iată, în substanţă, caracterizarea metodei statistice, a locului şi folosirii ei printre metodele psihologiei. Ea este indispensabilă pentru progresul psihologiei, fără a fi suficientă în efortul de constituire a adevărului psihologic. În ceea ce priveşte folosirea matematicii, rolul acesteia a devenit şi mai important prin folosirea ciberneticii şi introducerea noilor metode în psihologie, cum este cea a modelării. Roadele vor spori prin colaborarea mai strânsă şi mai organică între matematicieni şi psihologi"[40].

Este adevărat că psihologul nu se află în situaţia savantului lui Eminescu din *Scrisoarea I* care, „Precum Atlas în vechime sprijinea cerul pe umăr/ Aşa sprijină el lumea şi vecia într-un număr", dar analiza atentă a cîmpurilor de cercetare psihologică indică o prezenţă foarte vastă şi diferenţiată a exprimării matematice. Psihologia clinică şi psihologia artei rămân la exprimări „artistice", pe când psihologia experimentală, psihologia cognitivă, neuropsihologia cunosc formulări cantitative puternic exprimate, justificate de nevoile progresului ştiinţific. Nu există

38. Sintagmă preluată din Mircea Miclea, *Psihologie cognitivă*, Casa de editură „Gloria" S.R.L., Cluj-Napoca, 1994, p. 18.
39. *Ibidem.*
40. Vasile Pavelcu, *Drama psihologiei*, Editura Didactică şi Pedagogică, Bucureşti, 1972, p. 266.

nici un motiv pentru a considera că modelele elaborate în psihologie, în ciuda specificității lor, trebuie să fie de altă natură decât modelele elaborate în alte domenii ale cercetării științifice. Gheorghe Zapan realizează o modelare a activităților umane pe baza statisticii matematice într-o lucrare de erudiție a bibliografiei de specialitate[41].

O excelentă prezentare a modelelor psihologice de descriere a activității aparține unui grup de autori francezi, format din Claude Bonnet, Christiane Kekenbosch, Anh Nguyen Xuan, J.F. Richard și Charles Tijus[42]. În concepția lor, modelele în discuție au cinci caracteristici majore:

- Un model este o particularizare a unei teorii de funcționare cognitivă. El însoțește o teorie, dar se aplică la un tip de situație particulară și face ipoteze suplimentare în raport cu teoria a cărui prelungire este.
- Un model este o formalizare și, ca atare, se exprimă într-un limbaj formalizat. Acest gen de limbaj nu este ambiguu și permite calcule, adică deducții automate. O bună perioadă de timp limbajul logic și matematic erau singurele care permiteau calcule. Însă informatica a favorizat dezvoltarea altor limbaje formalizate care fac posibile, pe lângă calculele numerice, și calculele simbolice, adică deducții privind cunoștințele și semnificațiile.
- Un model este schematic, adică reprezintă o simplificare a realității care nu reține decât aspectele esențiale (este vorba despre aspectele care, în principiu, sunt suficiente pentru a prezice un comportament). Caracterul schematic este și o caracteristică a teoriilor; deci atât modelele, cât și teoriile revendică ideea după care aceste simplificări conservă trăsăturile esențiale ale realității prin raportare la tipul de comportament pe care îl explică.
- Un model furnizează predicții validabile prin date experimentale, deci el trebuie să „reziste" la o serie de provocări ale realității.
- Un model psihologic conține o descriere a situației la care el se aplică (nu se reține decât un anumit număr de elemente pertinente). Modelele conțin o descriere a mecanismelor de tratament informațional pentru acele tipuri de situații prin care se poate deduce, cu ajutorul unor proceduri de calcul automatic (datorită faptului că descrierea este realizată printr-un limbaj formalizat), comportamentul pe care-l va adopta un subiect ce se comportă precum modelul. Un astfel de procedeu de tratament informațional se numește simularea comportamentului.

Foarte importante în psihologie sunt modelele de măsură care au apărut din nevoia stabilirii relevanței informaționale a rezultatelor. Aceste modele pot stabili dacă datele pot fi raportate la una dintre scalele de măsură (nominală, ordinală, de log-interval, de interval, de raport și absolută).

41. Gheorghe Zapan, „Modelarea matematică a activităților umane și a proceselor evolutive cu specială aplicare în domeniile psihologiei și pedagogiei", în Zapan, Gheorghe, *Cunoașterea și aprecierea obiectivă a personalității*, Editura Științifică și Enciclopedică, București, 1984.
42. Studiul se intitulează „Mesure et modélisation", din volumul *Cours de Psychologie* (sous la direction de R. Ghiglione et J.F. Richard), Dunod, Paris, 1995, pp. 268-384.

Modelele de măsură sunt numeroase, dar nu toate prezintă aceeași importanță (o parte din ele se aseamănă și o parte sunt pe cale de experimentare, unele prezintă un accentuat interes euristic, iar altele un interes istoric). Prezentăm în continuare modelele folosite foarte frecvent[43]:

- Modelul legii judecăților comparative a lui Thurstone a apărut în anul 1927 și vizează construirea unei scale de interval plecând de la date adecvate, rezultate dintr-o mulțime de comparații a două câte două observații (care a servit mult timp ca referință la măsura unidimensională).
- Analiza structurilor latente, creată de Lazarsfeld, a apărut în anul 1959 și constituie un sistem coerent, dificil de disociat. O aplicație foarte curentă vizează analiza claselor latente a căror funcție este de a identifica la nivel latent clase de obiecte echivalente plecând de la variabile manifeste (nominale).
- Teoria clasică a scorurilor adevărate este destinată creării de indicatori compoziți, omogeni, plecând de la mai mulți itemi. Concepută în contextul testelor mentale, este cunoscută sub denumirea de teoria clasică a construcției testelor. Aceste teorii răspund la estimarea fidelității și se prelungesc în teoria generalizabilității.
- Modelele de răspuns la itemi sunt modele probabiliste legate între ele. Modelul cel mai simplu și cel mai utilizat a fost trasat de matematicianul danez Georg Rasch, în 1960; este vorba de fapt despre o versiune probabilistă a modelului lui Guttman. Funcția sa este de a ordona subiecți și itemi pe o dimensiune, plecând de la datele dihotomice. El ajunge la o măsură de interval și constituie o ameliorare netă în raport cu teoria clasică a scorurilor adevărate. Tinde să stabilească măsura itemilor independent de subiecții în funcție de care a fost stabilită, iar măsura subiecților rămâne independentă de itemii asupra cărora ei s-au pronunțat.
- Modelul de analiză factorială figurează printre modelele cele mai familiare psihometricienilor. Obiectivul comun este de a condensa relațiile între un mare număr de variabile manifeste, sub forma relației dintre câteva variabile latente (factori, componente) pe care le înlocuiesc. Acestea sunt modele multidimensionale care se conformează cel mai adesea exigențelor variabilității de interval.
- Modelul LISREL (prescurtare de la LInear Structural RELationship) este un sistem complex care înglobează un model de măsură (destinat studiului relațiilor intraconcept) și un model structural (destinat studiului relațiilor interconcept). Modelul de măsură este avantajos pus în aplicare pentru a urma un obiectiv identic cu cel al analizei factoriale, reclamă formularea explicită a unor ipoteze asupra organizării variabilelor și oferă, prin suplețea sa, posibilități variate de modelizare.

43. O analiză extinsă a unor astfel de modele se găsește în Paul Dickes; Jocelyne Tournois; André Flieller; Jean-Luc Kop, *La psychométrie. Théorie et méthode de la mesure*, PUF, Paris, 1994, pp. 139-247.

- Modelul de eşalonare multidimensională constituie o mare familie de modele, având funcţia de a reprezenta similaritatea sub forma distanţei geometrice într-un spaţiu multidimensional. Modelul cel mai utilizat tratează relaţiile ordinale şi, printr-o transformare ordinală, permite să se ajungă la o măsură de raport. Modelul nu introduce constrângeri puternice asupra datelor şi oferă numeroase variante.

Investigarea prin intermediul unui model se realizează în proporţii variabile. În multe cazuri, atât modelele, cât şi teoriile sunt implicite şi autorul testului nu insistă prea mult asupra părţii de fundamentare teoretică a instrumentului său. Tipul de solicitare pe care îl presupune sarcina psihologică trebuie să fie exprimat în forme concrete prin care să se realizeze exigenţa teoretică.

Pregătirea spre studiul mai amănunţit al unor astfel de modele necesită noţiuni din teoria probabilităţilor. Prezentarea în continuare a acestor noţiuni urmăreşte, pe de o parte, construirea unui ansamblu coerent şi, pe de altă parte, reducerea matematizării.

5.2. Construcţii axiomatice

Teoria probabilităţilor este o teorie matematică, adică o construcţie abstractă şi axiomatică. O teorie matematică tratează relaţii între concepte, noţiuni abstracte. Se introduc, fără a le defini, noţiuni numite concepte fundamentale (cum ar fi *punctul* din geometrie) şi se enunţă, fără a le demonstra, propoziţii, *axiomele* teoriei, care descriu proprietăţile respectivelor concepte fundamentale. Pornind de la conceptele fundamentale şi axiomele teoriei, se construieşte apoi întreaga teorie. Rezultatele obţinute prin raţionamente logice sunt numite *teoremele* teoriei.

Deşi, după schema foarte simplificată prezentată, se pot construi teorii diverse, dezvoltarea unei teorii cu aplicaţii la lumea realului este departe de a fi atât de arbitrară. Alegerea conceptelor fundamentale şi a axiomelor este motivată de dorinţa de a obţine o descriere idealizată – un model – a unor anumite fenomene observabile (direct sau prin consecinţele lor). În acest sens, conceptele fundamentale corespund, în general, noţiunilor experimentale, iar axiomele corespund proprietăţilor acestora. Pentru ca teoria să aibă credibilitate şi să se impună, este necesar ca unele dintre propoziţiile (teoremele) demonstrate în cadrul teoriei să poată fi confruntate cu fapte observate. Această confruntare, în verificarea propoziţiilor, cu lumea realului constituie baza acceptării altor propoziţii, a căror verificare este uneori foarte dificilă.

În procesul de modelare prezentat, trebuie să avem imaginea din figura 5.1, care prezintă, într-un mod schematizat, lumea realului şi modelul care încearcă organizarea acestei lumi.

Analiza figurii arată, pe lângă regularitatea modelului şi aparenta dezordine a realului, că sunt sugerate două situaţii:

1. în afara modelului rămân aspecte ale realului neprinse în model;
2. în interiorul modelului există părţi care nu au corespondent în realitatea care se modelează.

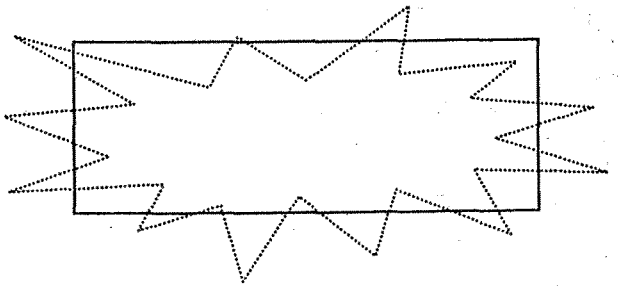


Figura 5.1. *Modelarea realului: modelul (linia plină), realul (linia punctată)*

Figura 5.1. nu vrea să introducă neîncrederea în posibilitatea modelării realului, ci să prevină încrederea absolută în rezultatele care nu pot fi verificate experimental. Totodată, se doreşte explicit introducerea unui grad de relativism al rezultatelor, mod de gândire propriu statisticii.

Teoria probabilităţilor este introdusă ca teoria matematică a fenomenelor aleatorii. Prin fenomen aleatoriu înţelegem un fenomen a cărui apariţie se află sub semnul hazardului, al întâmplării. Noţiunea contrară este aceea de fenomen determinat, adică fenomen a cărui apariţie este previzibilă prin cunoaşterea cauzalităţii sale. Uneori, caracterul aleatoriu este dat doar de necunoaşterea cauzei implicate, dacă există o asemenea cauză.

Comportamentul uman este condiţionat de un ansamblu complex de factori interni şi externi, fapt ce-l expune la fluctuaţii deseori imprevizibile. O aceeaşi persoană, în momente diferite, poate reacţiona diferit la situaţii identice, diferite persoane pot reacţiona diferit la situaţii identice etc.

Condiţionarea de context este frecvent întâlnită în lumea reacţiilor comportamentale umane. La rândul său, contextul poate să se realizeze pe plan subiectiv (motivaţie, afectivitate etc.), pe plan obiectiv sau poate rezulta din combinaţia celor două.

Fricosul poate reacţiona prin fugă la apariţia unei stimulări cu semnificaţie de pericol în 999 de cazuri dintr-o mie, dar poate deveni extrem de curajos în cazul 1000, dacă îi sunt puse în joc onoarea, imaginea socială sau când prin fugă ar periclita viaţa membrilor familiei etc.

Chiar şi reacţiile psihice mai simple au o evoluţie greu previzibilă. Viteza de reacţie la o stimulare de natură diferită (vizuală, sonoră, cutanată etc.) este imprevizibilă la nivel intraindividual.

Se pot face câteva precizări asupra viziunii probabiliste a legăturilor dintre fenomene. Această viziune nu este prezentă doar în matematică, disciplină care

formalizează de fapt un curent de gândire. Patrick Suppes creează chiar o metafizică probabilistă, care se bazează pe un număr de aserţiuni. Acestea sunt:

- Legile fundamentale ale fenomenelor naturale au în esenţă mai curând un caracter probabilist decât unul determinist.
- Concepţia noastră despre materie trebuie să conţină un element probabilist intrinsec.
- Cauzalitatea are un caracter probabilist, nu unul determinist; prin urmare, nu există incompatibilitate între aleatoriul din natură şi existenţa legilor cauzale valide.
- Certitudinea cunoaşterii – în sensul caracterului psihologic nemijlocit, în sensul adevărului logic sau în sensul preciziei totale a măsurătorilor – este irealizabilă.
- Colecţia teoriilor ştiinţifice trecute, prezente şi viitoare nu converge către un rezultat fixat inevitabil care să ofere, la limită, o cunoaştere completă a universului.
- Ştiinţele se caracterizează, în ceea ce priveşte limbajul, obiectul şi metoda, mai curând prin pluralism decât prin unitate.
- Însuşirea limbajului şi performanţa lingvistică, în aspectele lor fonologice, gramaticale, semantice şi prozodice, au caracter intrinsec probabilist.
- Teoria raţionalităţii are caracter intrinsec probabilist[44].

În psihologie, demersul inductiv este folosit cu prioritate şi îşi propune căutarea unor legături cauzale. Jean Piaget spune în această privinţă: „Inducţia participă într-o măsură sau alta la hazard: hazardul demersului subiectului în privinţa realului sau hazardul inerent realului însuşi"[45].

Ideea unei conexiuni cauzale între fenomenele naturii poate fi înlocuită prin ideea unei conexiuni de probabilitate. Dacă psihologul tinde să ajungă la legile adevărate acţionând asupra faptelor la nivelul său de observaţie, el descoperă de fapt doar legi de probabilitate. Nu înseamnă că se abandonează poziţiile deterministe. E. Brunswik face o remarcă generală: fiinţele vii nu primesc decât semnale ambigue şi incerte din mediul la care ele trebuie să se adapteze şi, ca atare, adaptarea se realizează doar pariind. Ele nu percep decât legi de probabilitate şi comportamentul lor nu poate fi reglat prin alte legi.

A cunoscut o largă dezvoltare studiul comportamentului subiecţilor plasaţi într-o situaţie care, la nivelul lor de observaţie, variază într-un mod aleatoriu.

Modelele matematice ale teoriei jocurilor, elaborate de von Neuman şi Morgenstern, au fost utilizate de psihologi şi economişti. Lazarsfeld a dorit ca aceste modele să nu fie construite plecând exclusiv de la situaţii prea simplificate şi ca ele să fie în măsură să se aplice la unele aspecte complexe ale vieţii cotidiene.

Este intrinsecă nesiguranţa în raport cu legile elaborate de psiholog datorită nivelului de observaţie, propriu psihologului, vizând comportamentele globale,

44. Patrick Suppes, *Metafizica probabilistă*, Editura Humanitas, Bucureşti, 1990, p. 72.
45. Jean Piaget, *Introduction à l'épistémologie génétique*, PUF, Paris, 1950, vol. II, p. 202.

molare. Psihologul nu va utiliza în fapt decât legi probabiliste. Acest lucru este recunoscut atât de cei care folosesc metode cantitative, cât şi de cei care nu recurg la ele. Cei care fac apel la metode cantitative se găsesc plasaţi într-o situaţie mai favorabilă decât cei care nu le folosesc. A recunoaşte incertitudinea unui prognostic este o simplă constatare, care poate fi utilă sau nu. Dacă se constată incertitudinea în termeni generali, deci calitativi, confirmarea sau infirmarea unei serii de prognosticuri particulare nu va permite ameliorarea metodei pe care aceste prognosticuri erau fondate. Incertitudinea cuantificată este una care „lucrează", în sensul în care cineva se referă la un organ îndeplinind o funcţie sau o sarcină. Este o incertitudine limitată şi diferenţială. Aceasta se poate regăsi în afirmaţii de genul: „Recunosc că sunt nesigur în ceea ce priveşte prognosticul de reuşită şcolară pe care îl realizez pe baza unui test, dar îmi impun limite între care estimaţiile mele trebuie să fluctueze. Cred că erorile mele sunt cu atât mai frecvente, cu cât sunt mai slabe, că voi face tot atâtea erori în plus câte erori în minus. În această privinţă mă pun în postura de a fi infirmat sau confirmat prin fapte şi în acest fel îmi asigur un mijloc de a îmbunătăţi modelul teoretic despre care am vorbit".

Se poate deci afirma că legile cauzale nu sunt accesibile psihologului decât sub forma legilor de probabilitate şi valoarea euristică a acestor legi este considerabil mărită atunci când ele pot să se exprime cantitativ.

5.3. Evenimente

Atunci când se studiază un fenomen aleatoriu, mai întâi trebuie să se stabilească lista rezultatelor posibile (uneori, faza respectivă coincide cu stabilirea scalei de măsură). Aceste noţiuni, studiul propriu-zis şi rezultatele posibile se regăsesc în teorie sub forma unei *probe* şi a *evenimentelor* asociate probei. „Probă" şi „eveniment" sunt concepte fundamentale ale teoriei probabilităţilor; drept urmare, nu sunt definite în funcţie de alte noţiuni. Ele idealizează, abstractizează noţiunile de experienţă, cercetare, investigaţie şi rezultatele observate în cercetare. De remarcat totuşi faptul că evenimentele fac parte dintr-o mulţime fixată. Astfel, în proba de aruncare a unui zar, dacă fixăm mulţimea evenimentelor la apariţia unei feţe numerotate a zarului, atunci apariţia feţei cu numărul 6 este un eveniment, dar spargerea zarului nu este un eveniment.

Mulţimea tuturor rezultatelor posibile este numită **mulţimea de bază** asociată probei şi se notează cu Ω.

- Dacă *mulţimea de bază este finită*, atunci elementele ei sunt numite **evenimente elementare**, iar orice mulţime de evenimente elementare (submulţime a mulţimii de bază) poartă numele de **eveniment.** Se poate spune astfel că este eveniment orice submulţime a mulţimii de bază, iar pentru mulţimea evenimentelor se utilizează notaţia $P(\Omega)$, notaţie consacrată pentru mulţimea părţilor lui Ω.

- Dacă *mulțimea de bază este infinită*, atunci mulțimea evenimentelor reprezintă, din considerente teoretice, o familie de submulțimi ale lui Ω care îndeplinesc condițiile:
  1) Ω aparține familiei;
  2) complementara oricărei mulțimi din familie aparține familiei;
  3) reuniunea oricător mulțimi din familie este tot o mulțime din familie.

Se poate observa că definiția mulțimii evenimentelor din cazul finit satisface și definiția cazului infinit.

Vom nota în continuare mulțimea evenimentelor cu A, iar evenimentele cu litere mari: A, B, C etc. În cazul finit reamintim că $A = P(\Omega)$.

Să vedem cum se prezintă datele cu care operăm în cercetarea psihologică sau cu care operează psihologul practician.

Procesul de măsurare în psihologie, care duce la crearea datelor, poate evidenția o ierarhie pronunțată a „forței" informației (problema scalelor de măsură). În cea mai mare parte, datele din cercetarea psihologică satisfac exigențele scalelor ordinale și de interval (după ce s-au acceptat, eventual, o serie de convenienșe). De asemenea, în mare parte, datele din domeniul psihologiei sunt discrete și formează mulțimi finite de eventualități. Amplitudinea distribuțiilor poate varia foarte mult în cazul unui test, de exemplu, dar în limite strict stabilite prin construcția probei.

În cadrul observației, avem un număr limitat de eventualități, evidențiate, mai ales în observația standardizată, prin așa-numitele *check-lists*.

Convorbirea și ancheta psihologică oferă, de asemenea, serii strict delimitate de eventualități.

Este interesant de evidențiat ce tipuri de variabile pot oferi suport, în cadrul unui experiment, pentru aplicații mai largi ale teoriei probabilităților.

Variabilele independente sunt acele variabile manipulate de experimentator. În marea majoritate a cazurilor, valorile lor formează un domeniu cantitativ strict determinat. Nu este exclusă randomizarea emisiei de stimuli, deci existența unui posibil domeniu al teoriei probabilităților.

În 1937, E.C. Tolman a recurs într-o conferință la expresia „variabilă intermediară". Pentru psihologul american, o teorie nu constituie decât „un sistem de variabile intermediare" (*intervening variable*), care sunt construcții „invocate" de către teoreticieni pentru a fragmenta într-un mod comod relația empirică stabilită de experimentator între variabilele independente și cele dependente, de exemplu între variațiile unui stimul și variațiile reacțiilor care îi succedă. În experiența asupra învățării la șoareci, el va explica relațiile dintre variabilele independente – condițiile de mediu (alimente distribuite, tipuri de stimuli, număr de încercări), diferențele individuale (ereditate, vârstă, antrenament prealabil) – și variabilele dependente (proporția alegerilor corecte în labirint) prin variabile intermediare precum nevoia, diferențierea, ipoteza. Nevoia apare în funcție de distribuție, iar diferențierea în funcție de stimuli.

Între evenimente se definesc operații, inspirate din teoria mulțimilor:

- **Egalitatea**: se spune că $A = B$ dacă apariția evenimentului A coincide cu apariția evenimentului B și reciproc. De exemplu, dacă fața 6 a unui zar este singura față colorată în verde, atunci evenimentul „6" este egal cu evenimentul „fața verde".
- **Negația**: se spune că evenimentul notat $\overline{A}$ este negația lui A dacă $\overline{A}$ apare atunci și numai atunci când nu apare A. De exemplu, la aruncarea unei monede cu fețele desemnate prin „cap" și „pajură", evenimentul „cap" este negația evenimentului „pajură". Se utilizează în acest caz și noțiunea de *evenimente contrare.*
- **Reuniunea**: spunem că evenimentul notat $A \cup B$ este reuniunea evenimentelor A și B dacă $A \cup B$ apare atunci și numai atunci când apare sau evenimentul A, sau evenimentul B, sau și A și B. Putem vorbi astfel de evenimentul „1 sau 2" la aruncarea unui zar ca fiind evenimentul care apare atunci când apare fața „1" sau fața „2", adică „1 sau 2" este egal cu „1" $\cup$ „2".
- **Intersecția**: spunem că evenimentul $A \cap B$ este intersecția evenimentelor A și B dacă $A \cap B$ apare atunci și numai atunci când apar concomitent A și B.
- **Includerea**: spunem că A este inclus în B, notat $A \subset B$, dacă evenimentul B apare de fiecare dată când are loc evenimentul A. Se poate spune că evenimentul A *implică* evenimentul B: producerea lui A duce la apariția simultană a lui B. De exemplu, evenimentul „august" implică evenimentul „vară". Diferența față de egalitatea evenimentelor constă în aceea că apariția evenimentului B nu produce, în mod obligatoriu, apariția evenimentului A (fapt cerut la definirea egalității).

În mulțimea evenimentelor se disting două evenimente cu caracter special:

- **evenimentul sigur** este evenimentul care se produce cu certitudine; se notează cu Ω. De exemplu, evenimentul „cap sau pajură" la aruncarea unei monede este evenimentul sigur.
- **evenimentul imposibil** este evenimentul care nu se produce niciodată; se notează cu ϕ. Este evident că evenimentul sigur este evenimentul contrar evenimentului imposibil și reciproc.

Se poate spune astfel că evenimentele sunt de trei categorii: sigure, imposibile și cu realizare „probabilă".

Două evenimente A și B se numesc **incompatibile** dacă intersecția lor este evenimentul imposibil, $A \cap B = \phi$, adică nu este posibil ca evenimentele A și B să se producă simultan.

Operațiile introduse între evenimente au proprietăți similare celor ale operațiilor între mulțimi. Nu sunt prezentate, deoarece nu sunt necesare nivelului de matematizare vizat în această lucrare. Utilizarea unora va fi însoțită de explicațiile necesare la momentul respectiv.

5.4. Probabilitate

În prezentarea şi prelucrarea datelor, măsurate pe diverse scale, noţiunea de frecvenţă relativă este o noţiune comună: se regăseşte la toate scalele, valorile obţinute pot fi utilizate pentru comparaţii etc. Practica arată că, atunci când o experienţă este repetată de un număr mare de ori, frecvenţa relativă a apariţiei unui fapt tinde să se stabilizeze. Astfel, de exemplu, la aruncarea unui zar, numărul de apariţii ale unei anumite feţe tinde să se stabilizeze la 1/6 (dacă zarul este perfect echilibrat). Numărând de câte ori apare un număr, să zicem 20, la extragerile sistemului Loto 6 din 49, într-o perioadă suficient de mare de timp, se va regăsi tendinţa de stabilizare a numărului de apariţii în jurul unei proporţii de 1/49 din totalitatea numerelor extrase.

Acest fenomen este formalizat în teoria probabilităţilor prin introducerea noţiunii de probabilitate.

Fie A mulţimea evenimentelor. Se numeşte **probabilitate** orice funcţie cu valori reale definită pe mulţimea evenimentelor, $P: A \to \Re$, şi care satisface următoarele:

a) probabilitatea oricărui eveniment este un număr nenegativ, adică mai mare sau egal cu zero: $(\forall)A, A \in A \Rightarrow P(A) \geq 0$;
b) probabilitatea evenimentului sigur este egală cu 1: $P(\Omega) = 1$;
c) probabilitatea oricărei reuniuni de evenimente incompatibile două câte două este egală cu suma probabilităţilor evenimentelor: dacă $A_i \cap A_j = \phi$, oricare ar fi indicii i şi j, atunci

$$P(A_1 \cup A_2 \cup \ldots \cup A_n \cup \ldots) = P(A_1) + P(A_2) + \ldots + P(A_n) + \ldots .$$

Se poate deduce din definiţia anterioară că, date fiind o mulţime de bază şi o mulţime de evenimente, se pot defini mai multe funcţii de probabilitate. Însă, o dată definită funcţia de probabilitate, toate rezultatele dezvoltate pentru probă vor fi determinate de respectiva alegere. Se spune în acest caz că tripleta (Ω, A, P) defineşte un **spaţiu de probabilitate.** După cum mulţimea de bază este sau nu finită, se poate vorbi de spaţiu de probabilitate finit sau spaţiu de probabilitate infinit. Spaţiul de probabilitate finit va fi notat şi prin tripleta $(\Omega, P(\Omega), P)$.

Prezentăm cele mai uzuale proprietăţi ale probabilităţii, demonstraţiile fiind, de regulă, imediate:

1) Dacă evenimentul A implică evenimentul B, atunci probabilitatea lui A este mai mică sau egală cu probabilitatea lui B: $A \subset B \Rightarrow P(A) \leq P(B)$.
2) Probabilitatea oricărui eveniment este mai mică sau egală cu 1: $P(A) \leq 1$. Rezultă, din definiţie, că probabilitatea oricărui eveniment este cuprinsă între 0 şi 1, inclusiv 0 şi 1.
3) Probabilitatea evenimentului contrar se obţine scăzând din 1 probabilitatea evenimentului direct: $P(\overline{A}) = 1 - P(A)$.

4) Probabilitatea evenimentului imposibil este 0: $P(\phi) = 0$.
5) Probabilitatea reuniunii a două evenimente oarecare se obţine scăzând din suma probabilităţilor individuale probabilitatea intersecţiei celor două evenimente: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. De remarcat că pentru două evenimente incompatibile se obţine relaţia c) din definiţie, scrisă pentru două evenimente.

Dacă spaţiul de probabilitate este finit şi notăm evenimentele elementare prin $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$, atunci rezultă că probabilitatea unui eveniment este suma probabilităţilor evenimentelor elementare care îl compun:

$$P(A) = \sum_{\omega_i \in A} P(\omega_i).$$

În cazul în care toate evenimentele elementare au aceeaşi probabilitate (situaţia este diferenţiată prin denumirea **evenimente echiprobabile**), adică:

$$P(\omega_i) = \frac{1}{n},$$

rezultă că

$$P(A) = \frac{n_A}{n},$$

unde n_A reprezintă numărul evenimentelor elementare care îl compun pe A. Relaţia ultimă este interpretată adesea astfel:

probabilitatea unui eveniment este egală cu raportul dintre numărul cazurilor favorabile şi numărul total de cazuri,

dar nu trebuie pierdut din vedere faptul că, pentru utilizarea acestui rezultat, evenimentele elementare trebuie să fie echiprobabile.

O situaţie des întâlnită în practică o reprezintă aceea în care cercetătorul este interesat de şansa de realizare a unui anumit eveniment, ştiind că a avut loc un alt eveniment. Această situaţie se poate interpreta ca fiind interacţiunea dintre evenimente sau vizează modul în care este modificată puterea de anticipare de către o cunoaştere prealabilă (experienţa personală). Teoria probabilităţilor reflectă acest lucru prin introducerea noţiunii de probabilitate condiţionată.

Se numeşte **probabilitatea evenimentului A condiţionată de evenimentul B** valoarea notată $P(A \mid B)$ şi definită prin raportul:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)},$$

unde P(B) este diferită de 0. Altfel spus, $P(A \mid B)$ reprezintă probabilitatea evenimentului A, ştiind că s-a realizat evenimentul B.

De exemplu, probabilitatea ca la aruncarea unui zar să apară „1" (evenimentul A) este de 1/6. Dacă se cunoaşte că a apărut o faţă cu un număr fără soţ

(evenimentul B), atunci probabilitatea lui A (condiționată de B) este 1/3. Într-adevăr, B = {„1", „3", „5"} și P(B) = 1/2 (jumătate dintre fețe sunt cu numere impare). Deoarece A este inclus în B, rezultă că intersecția dintre A și B este chiar A. Prin urmare:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)} = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}.$$

Evenimentele A și B se numesc **independente** dacă are loc

$$P(A \cap B) = P(A) \cdot P(B).$$

În acest caz, se verifică imediat că P(A | B) = P(A), adică realizarea evenimentului B nu modifică probabilitatea evenimentului A, ceea ce justifică denumirea de evenimente independente.

Ca o observație finală a acestei secțiuni, se poate menționa faptul că frecvențele relative verifică proprietățile probabilității, ceea ce susține introducerea funcției probabilitate ca model pentru faptul că frecvențele relative tind să se stabilizeze o dată cu mărirea numărului de observații.

5.5. Variabile aleatorii

Cu ajutorul variabilelor aleatorii se modelează în teoria probabilităților ceea ce în științele experimentale se înțelege prin caracteristică, variabilă studiată etc. Faptul poate fi intuit corect dacă se are în vedere, de exemplu, măsurarea înălțimii unui elev, într-o cercetare antropometrică. A spune că elevul X are înălțimea de 1,20 m înseamnă că mai întâi a fost selectat elevul X (*realizarea unui eveniment*) și apoi că operația de măsurare a avut ca rezultat valoarea de 1,20 m (*o valoare a fost atribuită evenimentului*). Deci, de cele mai multe ori, într-o cercetare, interesul cercetătorului este îndreptat către valorile care sunt atribuite diferitelor acte, fapte, evenimente. Prin intermediul variabilelor aleatorii se modelează exact acest proces experimental.

5.5.1. Cazul finit

Fie $(\Omega, P(\Omega), P)$ un spațiu de probabilitate finit, $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$. Se numește **variabilă aleatorie discretă** orice funcție cu valori reale definită pe Ω. Vom nota variabilele aleatorii cu X, Y, Z etc. și cu literele mici corespunzătoare valorile luate. Adeseori se va utiliza prescurtarea v.a. Prin urmare, orice funcție

$$X : \Omega \to \Re$$

este o variabilă aleatorie.

Se observă că, prin modul de definire, fiecărui eveniment elementar ω_i i se asociază un număr real $X(\omega_i)$. Deoarece aceste numere nu sunt, în mod necesar, distincte între ele, notăm cu $x_1, x_2, \ldots, x_k$, $k \le n$, valorile distincte luate de v.a. X. Este natural să luăm în considerare acum acele evenimente, nu neapărat elementare, pentru care v.a. X ia valorile distincte:

$$A_i = \{\omega \mid X(\omega) = x_i\},\ i = 1, 2, \ldots, k.$$

Deși construcția pare artificială, trebuie înțeles că realizarea evenimentului A_i are ca efect faptul că v.a. X ia valoarea x_i (altfel spus, în A_i sunt toate evenimentele elementare a căror realizare conduce la luarea de către v.a. X a valorii x_i). Este evident rezultatul că familia $A_1, A_2, \ldots, A_k$ formează o partiție a mulțimii de bază Ω, numită **partiția indusă** de v.a. X. Cu alte cuvinte, orice eveniment elementar aparține unui eveniment A_j unic și numai unuia singur.

Construcția prezentată în acest capitol a ajuns în stadiul reprezentat în figura 5.2, în care liniile pline arată legăturile stabilite deja. Se vede că oricărui eveniment A i se pune în corespondență probabilitatea P(A), iar, prin considerarea unei v.a. X, i se conferă o valoare X(A). Este natural să se stabilească și legătura nediscutată încă și figurată prin linia punctată. Se poate astfel considera că *o variabilă aleatorie X ia valori cu anumite probabilități*, și anume cu probabilitățile cu care se realizează evenimentele asociate acelor valori ale lui X.

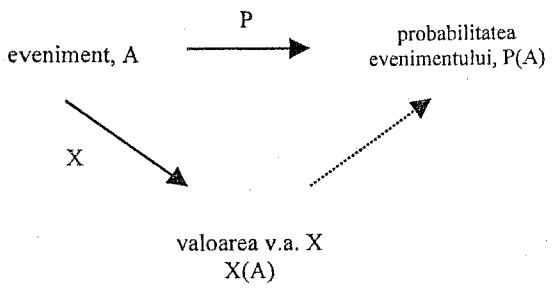


Figura 5.2. *Legăturile dintre noțiuni*

Acest lucru se formalizează prin luarea în considerare a **funcției de probabilitate** a variabilei aleatorii X, care se definește prin:

$$f : \{x_1, x_2, \ldots, x_k\} \to [0,1]$$
$$f(x_i) = P(A_i),\ i = 1, 2, \ldots, k.$$

Funcția de probabilitate a variabilei aleatorii discrete X se mai numește **legea de probabilitate** a v.a. X și se notează:

$$X:\begin{pmatrix} x_1 & x_2 & \dots & x_k \\ p_1 & p_2 & \dots & p_k \end{pmatrix},$$

unde $p_i = f(x_i)$ reprezintă probabilitatea cu care X ia valoarea x_i. O notaţie consacrată sugestivă este şi $p_i = P(X = x_i)$, $i = 1, 2, \dots, k$.

Din definiţiile şi proprietăţile probabilităţilor rezultă relaţiile:

$$f(x_i) \geq 0,\ i = 1, 2, \dots, k.$$

$$\sum_{i=1}^{k} f(x_i) = \sum_{i=1}^{k} P(X = x_i) = P(\Omega) = 1.$$

Semnificaţia lor este asemănătoare cu proprietăţile frecvenţelor relative de a fi nenegative şi de a avea suma egală cu 1.

Urmărind paralela dintre prelucrarea datelor empirice şi construcţia teoriei probabilităţilor, noţiunea de frecvenţă relativă cumulată se regăseşte în aceea de funcţie de repartiţie.

Numim **funcţie de repartiţie** a v.a. X funcţia reală de variabilă reală, $F: \Re \to \Re$, definită prin

$$F(x) = P(X \leq x),$$

unde prin $(X \leq x)$ s-a notat evenimentul $\{\varpi \mid X(\omega) \leq x\}$, adică reuniunea acelor evenimente elementare pentru care v.a. ia valori mai mici sau egale cu x.

Prin urmare, valoarea funcţiei de repartiţie într-un punct oarecare x este egală cu probabilitatea ca v.a. X să ia valori mai mici sau egale cu x. Se observă că, utilizând notaţiile anterioare,

$$P(X \leq x) = \sum_{x_i \leq x} P(A_i),$$

adică se însumează toate probabilităţile evenimentelor A_i care conduc la valori ale variabilei aleatorii X mai mici sau egale cu x. Se observă astfel asemănarea cu frecvenţele relative cumulate.

Două variabile aleatorii X şi Y definite pe o aceeaşi mulţime de bază Ω sunt independente dacă evenimentele partiţiilor induse de X şi Y sunt independente două câte două, adică pentru orice i şi j are loc

$$P(X = x_i \cap Y = y_j) = P(X = x_i) \times P(Y = y_j),$$

relaţie care exprimă faptul că variabilele aleatorii X şi Y iau valori independent una de alta. Dacă se vorbeşte de mai mult de două variabile independente înseamnă că ele sunt independente două câte două.

5.5.2. Cazul infinit

Fie un spaţiu de probabilitate (Ω, A, P). Se numeşte variabilă aleatorie orice funcţie reală definită pe Ω, $X: \Omega \to \Re$ şi care satisface condiţia:

$$\{\omega \mid X(\omega) \leq x\} \in A, \text{ oricare ar fi } x \in \Re.$$

Cu alte cuvinte, se cere, din considerente teoretice, ca *mulţimea evenimentelor elementare pentru care v.a. X are valori mai mici sau egale cu x, oricare ar fi x, să fie tot un eveniment.* Se poate considera că definiţia introdusă în cazul finit nu poate fi aplicată aici, deoarece împărţirea probabilităţii totale la un număr infinit de evenimente elementare produce o probabilitate egală cu zero pentru fiecare eveniment în parte. Din acest motiv, în cazul infinit, are sens să luăm în considerare probabilitatea ca v.a. X să aparţină la un anumit interval, şi nu probabilitatea să ia o anumită valoare.

Această interpretare nu este chiar atât de depărtată de practica experimentală, întrucât, de exemplu, şansa de a găsi un individ cu înălţimea de exact 1,80 m este aproape nulă dacă măsurarea se efectuează cu eroare zero. De obicei, printr-o asemenea valoare înţelegem un întreg interval de înălţimi, toate acelea care prin rotunjire la două zecimale devin 1,80 m.

Ca şi în cazul finit, numim **funcţie de repartiţie** a v.a. X funcţia reală de variabilă reală, $F: \Re \to \Re$, definită prin:

$$F(x) = P(X \leq x),$$

unde prin $(X \leq x)$ s-a notat evenimentul $\{\varpi \mid X(\omega) \leq x\}$ (aici apare, de exemplu, necesitatea condiţiei din definiţia unei v.a.), adică reuniunea acelor evenimente elementare (fiind o infinitate, ele nu se pot decât cel mult enumera) pentru care v.a. ia valori mai mici sau egale cu x.

Funcţia de repartiţie este absolut continuă dacă există o funcţie reală, $f: \Re \to \Re$, astfel încât

$$F(x) = \int_{-\infty}^{x} f(u)\,du,$$

cu interpretarea uzuală de mărime a ariei de sub graficul funcţiei f, din figura 5.3.

Funcţia f, dacă există, se numeşte **densitate de probabilitate** a v.a. X şi are proprietăţile:

$$f(x) \geq 0, \text{ oricare ar fi } x,$$

$$\int_{-\infty}^{+\infty} f(u)\,du = 1.$$

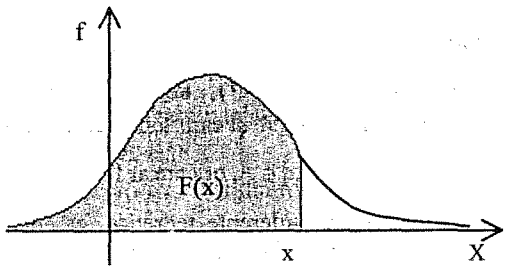


Figura 5.3. *Interpretarea geometrică a funcției de repartiție*

Se regăsesc astfel proprietățile legii de probabilitate pentru o variabilă aleatorie discretă.

Proprietățile definitorii ale funcției de repartiție sunt:

1) $0 \le F(x) \le 1$;

2) funcția F este nedescrescătoare (dacă $x_1 < x_2$, atunci $F(x_1) \le F(x_2)$);

3) $\lim_{x \to -\infty} F(x) = 0$; $\lim_{x \to +\infty} F(x) = 1$;

Dacă F este funcție de repartiție absolut continuă a variabilei aleatorii X, atunci are loc și proprietatea:

4) $P(x_1 \le X < x_2) = F(x_2) - F(x_1)$,

care se aplică la calculul probabilităților atunci când se cunoaște funcția de repartiție.

De remarcat că proprietățile nu aduc ceva principial nou, deoarece 1) rezultă din domeniul unei probabilități, 2) reflectă faptul că probabilitatea unui eveniment nu poate să scadă dacă se reunește un alt eveniment și 3) reflectă, de fapt, probabilitățile pentru evenimentul imposibil și, respectiv, evenimentul sigur.

Funcția de repartiție conține toată informația necesară pentru calcularea probabilităților cu care variabila aleatorie ia valori și pentru acest lucru va fi utilizată în ceea ce ne interesează.

Reprezentarea grafică a unei funcții de repartiție are formă distinctă după cum funcția este discretă sau continuă. Aspectul general este însă același datorită proprietăților generale. Figura 5.4 conține graficele unor funcții de repartiție.

Se observă forma în trepte a funcției de repartiție pentru o v.a. discretă, corespunzătoare trecerii de la o valoare la alta. Suma tuturor salturilor, marcate pe grafic cu săgeți, este egală cu 1.

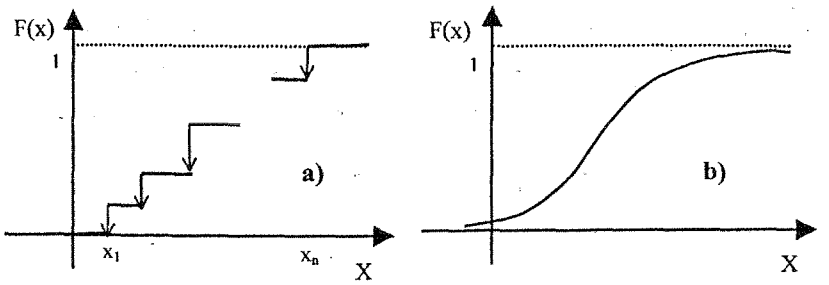


Figura 5.4. *Funcția de repartiție – forma generală a graficului: a) v.a. discretă; b) v.a. continuă.*

5.6. Valori tipice ale variabilelor aleatorii

Caracterizarea valorilor unei variabile aleatorii, atât pentru studiul ei, cât și în vederea comparării cu altă variabilă aleatorie, trebuie să fie suficient de simplă și flexibilă. Acest deziderat este atins prin introducerea unor *rezumate* sub formă de valori tipice. Prin **valoare tipică** înțelegem o valoare calculată pe baza variabilei aleatorii și a probabilităților asociate, valoare care cuprinde într-o formă sintetică un aspect de interes al comportamentului variabilei aleatorii respective. Sunt urmărite, în special, aspecte legate de gruparea valorilor, împrăștierea lor și trăsături caracteristice ale reprezentărilor grafice ale repartițiilor.

Prezentarea sumară a principalelor valori tipice este dată atât pentru întregirea imaginii legate de variabilele aleatorii, cât și pentru apropierea de noțiunile din capitolele precedente.

5.6.1. Valoarea medie

Fiind dată o v.a. X, se definește **valoarea medie**, notată M(X) sau μ, prin:

- dacă X este discretă, atunci:

$$M(X) = \sum_i x_i p_i;$$

- dacă X este continuă, atunci:

$$M(X) = \int_{-\infty}^{+\infty} x f(x)\, dx$$

Se observă că definiția este similară mediei aritmetice; cazul discret este media aritmetică ponderată a valorilor, iar cazul continuu – generalizarea primului caz.

Dintre proprietăţile valorii medii amintim:

- media unei sume finite de variabile aleatorii este egală cu suma mediilor:

$$M(X_1+X_2+\ldots+X_n)=M(X_1)+M(X_2)+\ldots+M(X_n);$$

- media produsului unor variabile aleatorii reciproc independente este egală cu produsul mediilor:

$$M(X_1X_2\ldots X_n)=M(X_1)\cdot M(X_2)\cdot\ldots\cdot M(X_n).$$

În cazul particular când o variabilă este constantă, au loc relaţiile:

$$M(c+X)=c+M(X),$$
$$M(c\times X)=c\times M(X),$$

care arată regulile de calcul utile în cazul în care se procedează la transformări prealabile ale variabilelor.

5.6.2. Quantila de ordin α

Quantila de ordin α se defineşte ca fiind acea valoare x_α a variabilei aleatorii X pentru care are loc

$$F(x_\alpha)=P(X<x_\alpha)=\alpha.$$

Quantila de ordin 1/2 se numeşte **mediană** şi este notată Me. Deci

$$P(X<Me)=P(X>Me)=\frac{1}{2}.$$

Se observă că, prin luarea în considerare a quantilelor de diferite ordine, se poate aprecia modul de dispunere a valorilor celor mai frecvente (cu probabilităţi mai mari de a fi luate).

5.6.3. Valoarea mod

Valoarea mod sau **modul**, **moda**, **valoarea modală**, notată Mo, este cea mai probabilă valoare a variabilei aleatorii.

Dacă v.a. X este discretă, atunci valoarea mod este valoarea cu cea mai mare probabilitate. Dacă v.a. X este continuă, atunci Mo este valoarea pentru care densitatea de probabilitate este maximă.

O v.a. cu o singură valoare mod este numită **unimodală**; dacă sunt mai multe valori modale, atunci se pot utiliza termenii: bimodală, plurimodală.

> **Observaţie.** Valoarea medie, mediana şi valoarea mod sunt utilizate pentru a analiza tendinţa centrală de grupare a valorilor unei variabile aleatorii. Se consideră de obicei că valorile se acumulează acolo unde probabilitatea este mai mare sau se interpretează că valorile care corespund unor probabilităţi mai mari sunt luate mai des de către v.a. (în acest fel, are loc o acumulare a valorilor mai probabile).

5.6.4. Momentul de ordin r

Prin **momentul de ordin r**, notat M_r, al unei v.a. X se înţelege valoarea medie a variabilei X^r, adică:

- dacă v.a. X este discretă:

$$M_r=M(X^r)=\sum_i p_i x_i^r;$$

- dacă v.a. X este continuă:

$$M_r=M(X^r)=\int_{-\infty}^{+\infty} x^r f(x)\,dx.$$

Se observă că valoarea medie este momentul de ordin 1.

Prin ridicarea la puterea r a valorilor unei v.a. X se realizează de fapt o ponderare mai mare a valorilor extreme. Dacă puterea este pară, atunci dispar diferenţele de semn; dacă puterea este impară, diferenţele de semn se păstrează. Interpretarea momentelor de diferite ordine se bazează pe această observaţie.

5.6.5. Momentul centrat de ordin r

Prin **momentul centrat de ordinul r**, notat μ_r, al unei v.a. X se înţelege momentul de ordin r al abaterii variabilei aleatorii X de la media ei:

$$\mu_r=M([X-M(X)]^r),$$

care se înţelege în succesiunea: – calculul valorii medii; – calculul abaterii de la medie; – ridicarea la puterea r; – calculul valorii medii pentru variabila obţinută la pasul anterior.

Observaţia făcută la momentul de ordinul r, cu privire la motivaţia ridicării la puterea r, se aplică şi aici, dar cu precizarea că utilizarea abaterii de la medie conduce la ponderarea mai mare a valorilor mai depărtate de la medie.

Dintre momentele centrate se evidenţiază momentul centrat de ordinul 2, numit dispersie şi care are un rol central în descrierea unei variabile aleatorii.

5.6.6. Dispersia

Dispersia unei v.a. X, notată $D^2(X)$ sau σ^2, este momentul centrat de ordinul 2, adică media pătratului abaterii [X–M(X)]:

- dacă X este discretă:

$$D^2(X)=\sum_i p_i[x_i-M(X)]^2;$$

- dacă v.a. X este continuă:

$$D^2(X)=\int_{-\infty}^{+\infty}[x-M(X)]^2 f(x)\,dx.$$

Dintre proprietăţile dispersiei amintim:

- dispersia unei constante este egală cu zero:

$$D^2(c) = 0;$$

- dispersia produsului dintre o constantă şi o v.a. este egală cu produsul dintre pătratul constantei şi dispersia v.a.:

$$D^2(cX) = c^2 D^2(X);$$

- dispersia unei sume finite de variabile aleatorii este egală cu suma dispersiilor:

$$D^2(X_1+X_2+\ldots+X_n) = D^2(X_1) + D^2(X_2) + \ldots + D^2(X_n).$$

Aplicând aceste proprietăţi, se obţin imediat alte două proprietăţi interesante (c este o constantă):

$$D^2(c+X) = D^2(X) \text{ şi}$$

$$D^2(X_1-X_2) = D^2(X_1) + D^2(X_2),$$

care pot oferi simplificări ale calculelor prin transformări prealabile ale variabilelor aleatorii.

Interpretând dispersia ca pe o măsură a împrăştierii valorilor cu probabilităţi suficient de mari (încât realizarea lor să fie plauzibilă), proprietăţile anterioare sunt instructive prin aceea că arată cum se împrăştie valorile rezultate prin operaţii între variabile aleatorii. De exemplu, diferenţa a două v.a. este tot atât de împrăştiată ca şi suma lor etc.

5.6.7. Abaterea standard

Abaterea standard sau **abaterea medie pătratică**, notată D(X) sau σ, se defineşte ca fiind rădăcina pătrată din dispersie:

$$D(X) = \sqrt{D^2(X)} = \sqrt{\sigma^2} = \sigma$$

unde ultima relaţie trebuie gândită ca succesiune de operaţii: mai întâi se calculează dispersia şi abia apoi se extrage rădăcina pătrată pentru a obţine abaterea standard.

Introducerea abaterii standard are avantajul că aceasta se exprimă în aceleaşi unităţi ca şi valorile variabilei aleatorii (dispersia se exprimă în unităţi pătrate).

5.6.8. Coeficientul de variaţie

Întrucât evaluarea împrăştierii prin intermediul dispersiei sau al abaterii standard nu poate fi utilizată în comparaţii (depind de ordinul de mărime al valorilor frecvente ale variabilei aleatorii), se introduce *coeficientul de variaţie*, notat V, care produce o mărime adimensională (nu depinde de unitatea de măsură):

$$V = \frac{\sigma}{|M(X)|}$$

şi care se exprimă uneori sub formă procentuală:

$$V = \frac{\sigma}{|M(X)|} \cdot 100(\%).$$

Interpretarea este evidentă prin raportarea abaterii standard la medie.

Observaţie. Dispersia, abaterea standard, coeficientul de variaţie şi momentele de ordin superior sunt utilizate pentru a aprecia împrăştierea valorilor unei variabile aleatorii. La fel ca la aprecierea tendinţei centrale a valorilor, se va considera că valorile cele mai probabile sunt mai frecvente. Rezultă astfel că o împrăştiere mai mare reflectă existenţa mai multor valori frecvente, pe când o împrăştiere mai mică reflectă mai puţine valori cu frecvenţe relativ mari.

5.6.9. Coeficientul de asimetrie

Pentru aprecierea gradului de simetrie a unei repartiţii se utilizează momentul centrat de ordinul al treilea, μ_3, care ponderează suplimentar valorile mai depărtate de medie (evident, ponderarea este importantă doar dacă valoarea respectivă are o probabilitate suficient de mare de realizare), dar conservă şi semnul abaterii de la medie. Rezultă astfel că valori negative ale momentului centrat de ordinul 3 reflectă repartiţii asimetrice spre stânga, iar valori pozitive reflectă repartiţii asimetrice spre dreapta.

Valoarea de comparaţie a indicatorului μ_3 este redusă, deoarece are acelaşi ordin de mărime cu valorile variabilei aleatorii pentru care se calculează. Standardizarea lui, pentru a permite comparaţii între repartiţii, se efectuează prin introducerea **coeficientului de asimetrie** (*skewness*) a repartiţiei variabilei aleatorii X:

$$\beta_1 = \frac{\mu_3^2}{\left(\sigma^2\right)^3},$$

unde σ^2 este dispersia v.a. X.

Întrucât formula anterioară produce întotdeauna o valoare pozitivă, se utilizează şi coeficientul:

$$\beta_1' = \frac{\mu_3}{\left(\sigma^2\right)^{3/2}}$$

care păstrează semnul, deci permite o apreciere a sensului asimetriei.

Cu cât β_1 (sau β_1') are o valoare apropiată de zero, cu atât repartiţia v.a. X este mai apropiată de o repartiţie simetrică.

Este evident că simetria unei repartiţii se reflectă în forma simetrică a graficului densităţii de probabilitate sau al funcţiei de probabilitate. Din acest motiv se poate considera că indicatorul se referă la forma graficului densităţii (funcţiei) de probabilitate.

5.6.10. Coeficientul de boltire. Excesul

Al doilea atribut important (primul fiind simetria) în aprecierea formei unei repartiţii *unimodale* este gradul de înălţare în apropierea valorii mod. Acest fapt poate fi măsurat prin panta curbei densităţii de probabilitate şi este definit prin:

$$\beta_2 = \frac{\mu_4}{\left(\sigma^2\right)^2},$$

unde μ_4 este momentul centrat de ordinul 4. Cantitatea β_2 se numeşte **coeficient de boltire** (*kurtosis*) sau **coeficient de aplatizare**. Se observă forma standardizată a acestui coeficient, ceea ce permite utilizarea lui în comparaţii între variabile aleatorii.

O valoare mai mare a coeficientului de boltire arată o supraînălţare în apropierea valorii mod, deci reflectă faptul că valorile respective au probabilităţi de realizare mult mai mari decât valorile mai depărtate.

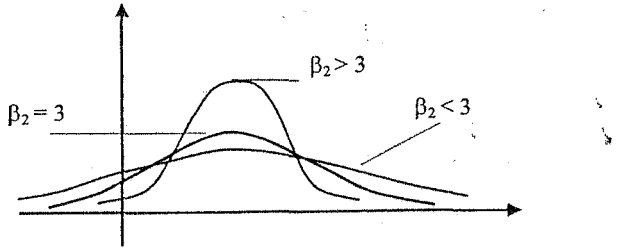


Figura 5.5. *Tipurile de repartiţii, după coeficientul de boltire*

Utilizarea coeficientului β_2 este realizată uneori prin raportare la boltirea repartiţiei normale (vezi paragraful următor), boltire egală cu 3. Se introduce astfel **excesul** unei v.a. X prin

$$E = \beta_2 - 3.$$

În funcţie de valoarea excesului, repartiţia este numită:

- *platicurtică*, dacă E < 0 (sau $\beta_2 < 3$);
- *mezocurtică*, dacă E = 0 (sau $\beta_2 = 3$);
- *leptocurtică*, dacă E > 0 (sau $\beta_2 > 3$).

Figura 5.5 ilustrează situaţiile menţionate luând în considerare diferite densităţi de probabilitate.

Cunoaşterea tipului de repartiţie permite aprecieri asupra grupării valorilor cu probabilităţi mai mari de realizare.

> **Observaţie.** Coeficientul de asimetrie şi coeficientul de boltire sunt utilizaţi pentru aprecierea formei unei repartiţii. De obicei, noţiunile sunt extinse asupra formei graficului densităţii de probabilitate.

5.7. Repartiţii teoretice

În această secţiune se prezintă repartiţiile teoretice care sunt implicate frecvent în analiza datelor experimentale. Trebuie precizat de la început faptul că utilizarea unui model teoretic nu rezolvă toate problemele experimentale şi are, de multe ori, un grad scăzut de aproximare, încât rezultatele obţinute sunt doar orientative. Acest fapt reflectă doar limitele cunoaşterii, şi nu faptul că o modelare nu este posibilă în domeniul studiat. De cele mai multe ori, nici nu există altă alternativă de prelucrare, astfel încât, chiar orientative, rezultatele sunt utile.

5.7.1. Repartiţia normală

Această repartiţie are un rol esenţial în teoria probabilităţilor şi în statistica matematică, atât din considerente teoretice, cât şi practice (nu în ultimul rând, pentru uşurinţa aplicării). Teoretic, repartiţia normală reprezintă o repartiţie-limită către care tind, în anumite condiţii, celelalte repartiţii. Practic, atunci când nu se cunoaşte repartiţia unui anumit fenomen, se poate considera, cu respectarea unui număr minim de condiţii, că fenomenul urmează o repartiţie normală. Această presupunere poate fi verificată, iar erorile produse în urma unei asemenea aproximări nu vor fi, în general, catastrofale. Este însă la fel de adevărat că exagerările de utilizare a repartiţiei normale, ca model teoretic, pot conduce la erori majore.

Prin definiţie, o variabilă continuă X are o **repartiţie normală** (numită şi **repartiţie Gauss-Laplace**), dacă funcţia de repartiţie este dată de:

$$F(x) = P(X < x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(t-\mu)^2}{2^2}} dt, \quad x \in \Re, \mu \in \Re, \sigma > 0,$$

unde μ şi σ sunt parametrii funcţiei de repartiţie.

Funcţia de repartiţie normală se va nota prin $N(\mu; \sigma^2)$, iar faptul că v.a. X este repartizată normal cu parametrii μ şi σ se notează $X \sim N(\mu; \sigma^2)$.

Prin definiţie, densitatea de probabilitate este dată de

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(x-\mu)^2}{2\sigma^2}}.$$

Se demonstrează că parametrii repartiției au semnificația unor valori tipice, și anume

$$M(X) = Me(X) = Mo(X) = \mu$$

$$D^2(X) = \sigma^2,$$

motiv pentru care se poate vorbi de repartiția normală cu media μ și dispersia σ^2, ceea ce determină complet repartiția.

Graficul funcției densitate de probabilitate f se numește **curba normală** sau **clopotul lui Gauss** și trebuie menționate următoarele proprietăți (demonstrația unora rezultă direct din definiții, altele necesită tehnici din analiza matematică):

- este simetrică față de dreapta $x = \mu$, iar ramurile curbei tind asimptotic la axa absciselor;
- este unimodală, valoarea maximă $\frac{1}{\sigma\sqrt{2\pi}}$ fiind atinsă pentru $x = \mu$;
- în vecinătatea punctului maxim, convexitatea curbei este spre exterior, iar în punctele de abscise $\mu + \sigma$ și $\mu - \sigma$ convexitatea curbei se schimbă (cele două puncte sunt puncte de inflexiune ale graficului);
- modificarea parametrului μ (păstrând σ constant) produce o translație a curbei fără modificarea formei; modificarea lui σ (păstrând μ constant) duce la o „ascuțire" a curbei când σ se micșorează și la o aplatizare a curbei când σ crește; situația corespunde ideii comune despre împrăștiere.

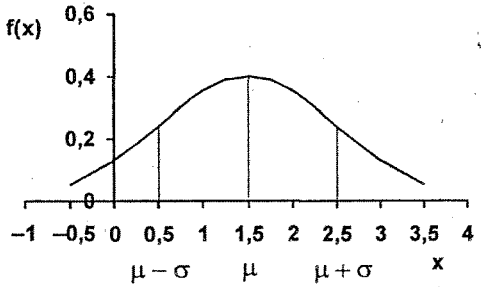


Figura 5.6. *Curba normală cu μ=1,5 și σ=1*

În figura 5.6 este reprezentată o curbă normală cu parametrii μ=1,5 și σ=1. Se observă punctul maxim de la abscisa 1,5 (= media) și punctele de inflexiune de la abscisele 0,5 și 2,5 (= $\mu - \sigma$ și $\mu + \sigma$).

În continuare sunt enunțate, fără demonstrație, principalele proprietăți ale repartiției normale:

a) Dacă X este o v.a. cu repartiție $N(\mu; \sigma^2)$, atunci momentele centrate sunt date de:

$$\mu_{2r} = \frac{(2r)!}{2^r \cdot r!}\sigma^{2r},$$

$$\mu_{2r+1} = 0.$$

Se observă că momentele centrate de ordin impar sunt egale cu zero datorită simetriei repartiției. Rezultă de aici că, pentru repartiția normală, coeficienții de asimetrie și boltire sunt:

$$\beta_1 = 0; \quad \beta_2 = 3.$$

b) Suma unor variabile independente repartizate normal este tot o variabilă aleatorie repartizată normal.

Dacă $X_1, X_2, \ldots, X_k$ sunt v.a. independente având repartiții normale cu media μ_i și dispersia σ_i^2, atunci suma $X = X_1 + \ldots + X_k$ este o v.a. repartizată normal cu parametrii

$$\mu = \sum_i \mu_i \text{ și } \sigma^2 = \sum_i \sigma_i^2 .$$

c) Proprietatea b) poate fi generalizată pentru o combinație liniară finită de v.a. independente.

Dacă $X_1, X_2, \ldots, X_k$ sunt v.a. independente având repartiții normale cu media μ_i și dispersia σ_i^2, atunci suma $X = \alpha_1 X_1 + \ldots + \alpha_k X_k$ este o v.a. repartizată normal cu parametrii

$$\mu = \sum_i \alpha_i \mu_i \text{ și } \sigma^2 = \sum_i \alpha_i^2 \sigma_i^2 .$$

Aplicând ultima proprietate, rezultă că prin transformarea de variabilă definită de

$$Z = \frac{X - \mu}{\sigma}$$

se obține o v.a. repartizată normal cu media 0 și dispersia egală cu 1, adică N(0;1). Faptul că orice variabilă repartizată normal poate fi transformată într-o v.a. repartizată N(0;1) asigură un rol important acestei repartiții particulare.

Pentru o astfel de situație se mai folosesc și denumirile **repartiție normală redusă, repartiție normală normată** sau **repartiție normală standard.** O v.a. repartizată N(0;1) este notată, în mod uzual, prin Z și este desemnată prin variabilă Z, variabilă normală redusă etc.

Graficul funcției densitate de probabilitate este reprezentat în figura 5.7. Graficul este simetric față de axa ordonatelor, yy', iar punctele de inflexiune au abscisele $z = -1$ și $z = +1$. În origine, funcția își atinge valoarea maximă $f(0) = 0{,}3989$.

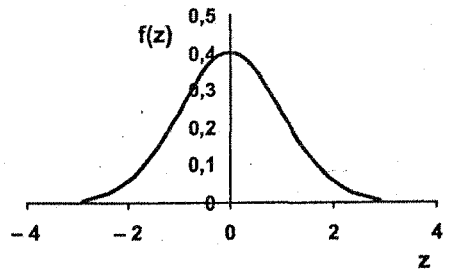


Figura 5.7. *Densitatea de probabilitate normală redusă*

Valorile densității de probabilitate, f(z), sunt tabelate în cărțile de statistică și sunt accesibile în orice program de calculator dedicat prelucrărilor statistice. Prin intermediul transformării Z, se pot calcula diferitele valori ale oricărei densități de probabilitate normale, indiferent de parametrii ei.

Mai importantă pentru aplicații, deoarece permite calculul probabilităților asociate diferitelor valori, este funcția de repartiție normală redusă:

$$F(z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-\frac{u^2}{2}}du.$$

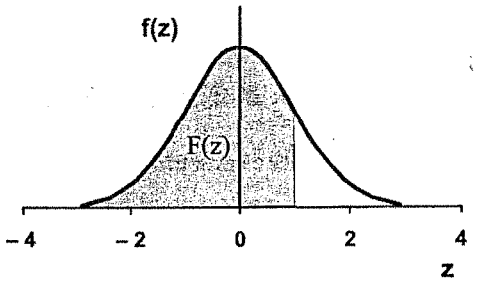


Figura 5.8. *Funcția de repartiție normală redusă*

Graficul acestei funcții este simetric față de axa yy' și are, potrivit interpretării uzuale din teoria funcțiilor, semnificația ariei de sub graficul funcției densitate de probabilitate. Deoarece, datorită simetriei, F(0) = 0,5, se introduce funcția:

$$\Phi(z)=F(z)-\frac{1}{2},\ z\geq 0,$$

care are semnificația ariei de sub ramura pozitivă a densității de probabilitate. Figurile 5.8 și 5.9 reprezintă semnificațiile geometrice ale funcțiilor F și Φ.

Atât valorile funcției F, cât și valorile funcției Φ sunt tabelate în lucrările de statistică.

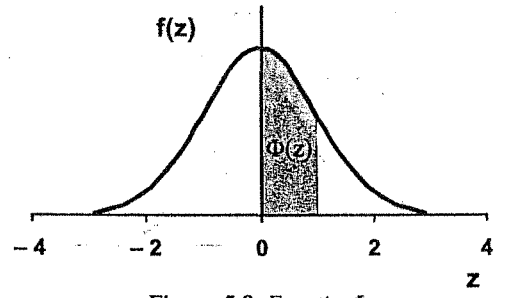


Figura 5.9. *Funcția Φ*

Un tabel tipic este organizat în 10 coloane și un număr variabile de linii. Valoarea aflată la intersecția unei linii cu o coloană este valoarea funcției F (sau Φ) corespunzătoare lui z definit prin:

- partea întreagă și prima zecimală este eticheta liniei;
- a doua zecimală este eticheta coloanei.

Astfel, valoarea F(1,24) se va afla la intersecția liniei corespunzătoare lui 1,2 (partea întreagă și prima zecimală) cu coloana corespunzătoare lui 4 (a doua zecimală; unii autori notează chiar coloana cu 0,04). Începutul tabelului este arătat în continuare:

z	0	1	2	3	4	5	6	7	8	9
0,0	0,5000	0,5040	0,5080	0,5120	...	...				
0,1	0,5398	0,5438	...	...						

Deoarece prima casetă conține valoarea 0,500, rezultă că este tabelată funcția F. Tabelul poate fi utilizat și pentru aflarea funcției inverse: dată fiind o valoare a funcției F, fie α această valoare, să se determine valoarea z pentru care F(z) = α. Din definiția funcției de repartiție rezultă atunci:

$$\alpha = F(z) = P(Z < z).$$

Definiția quantilelor unei variabile aleatorii arată în acest caz că valoarea z astfel determinată este quantila de ordin α a variabilei normale reduse Z. Se notează această valoare prin z_α. Reamintim interpretarea potrivit căreia valoarea z_α este acea valoare care depășește (100α)% dintre valorile v.a. Z. Deoarece este foarte frecventă interpretarea procentajului de valori care depășesc o limită dată, quantilele sunt notate și prin $z_{1-\alpha}$, α reprezentând proporția de valori ce depășesc quantila. Trebuie astfel reținută, pentru o bună interpretare, figura 5.10, care arată cum descompune quantila z_α aria totală de sub graficul densității de probabilitate în două arii de mărime $1-\alpha$ și, respectiv, α. Fiecare arie are semnificația probabilității cu care v.a. normală redusă Z ia valori în porțiunea

respectivă. Este ca şi cum s-ar ţinti, la întâmplare, un punct de sub graficul funcţiei f şi am avea şansa de $100(1-\alpha)\%$ de a nimeri zona haşurată (din stânga) şi şansa e $100\alpha\%$ de a nimeri zona nehaşurată (din dreapta).

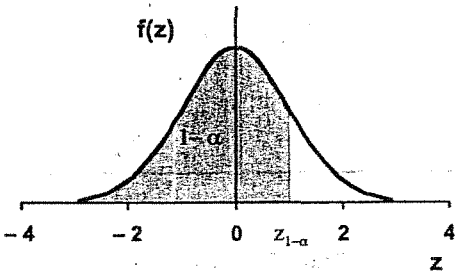


Figura 5.10. *Interpretarea uzuală a quantilelor*

În aplicaţii sunt utile următoarele relaţii pentru o variabila aleatorie X oarecare, repartizată normal, $X \sim N(\mu;\sigma^2)$:

a) calculul probabilităţii pentru o v.a. X oarecare:

$$F(x) = P(X < x) = P(Z < \frac{x<\mu}{\sigma}) = F(z), \text{ unde } z = \frac{x-\mu}{\sigma};$$

b) calculul probabilităţii ca o v.a. X să ia valori între două valori date:

$$P(x_1 < X < x_2) = P(X < x_2) - P(X < x_1) = F(z_2) - F(z_1),$$

unde z_1 şi z_2 sunt transformatele Z ale valorilor x_1 şi x_2;

c) calculul probabilităţilor ca o v.a. X să ia valori raportate la medie şi abatere standard:

$$P(X < \mu + k\sigma) = P(\frac{X-\mu}{\sigma} < k) = P(Z < k) = F(k)$$
$$P(X > \mu + k\sigma) = P(Z > k) = P(Z < -k) = F(-k) = 1 - F(k)$$
$$P(|X-\mu| < k\sigma) = P(-k\sigma < X-\mu < k\sigma) = P(-k < Z < k) = 2F(k) - 1,$$

unde k este o constantă pozitivă.

Ultima relaţie este utilă prin aceea că exprimă probabilitatea cu care valorile v.a. X se abat de la medie cu mai puţin de k abateri standard. Cantitatea $2F(k)-1$, exprimată eventual procentual, este numită uneori **siguranţă statistică.** Poate fi interpretată ca o măsură a încrederii prin faptul că v.a. X ia valori în intervalul $(\mu - k\sigma, \mu + k\sigma)$.

5.7.2. Repartiţia log-normală

O v.a. Y are o repartiţie log-normală cu parametrii μ şi σ^2, dacă variabila transformată prin logaritmare, $X = \ln Y$, are o repartiţie normală cu parametrii respectivi.

Repartiţia log-normală poate fi utilizată pentru modelarea unor caracteristici ale căror valori diferă foarte mult ca ordine de mărime. Prin logaritmare se atenuează diferenţele şi se poate obţine o aproximare mai bună sub forma simetrică a distribuţiei normale.

Are loc

$$P(y_1 < Y < y_2) = P(\ln y_1 < X < \ln y_2) = F(\ln y_2) - F(\ln y_1),$$

unde F este repartiţia normală redusă.

5.7.3. Repartiţia χ^2

Cunoscută şi sub numele de repartiţia CHI-2, sub influenţa literaturii în limba engleză şi a programelor de calculator, repartiţia χ^2 este utilizată mai mult în testarea ipotezelor statistice decât ca model statistic. Din acest motiv vom defini repartiţia χ^2 drept repartiţia unei v.a. care se obţine ca suma pătratelor unor v.a. independente repartizate normal standard, adică:

dacă $Z_1, Z_2, \ldots, Z_\nu$ sunt variabile aleatorii independente, fiecare cu repartiţia N(0;1), atunci variabila aleatorie

$$\chi^2 = Z_1^2 + Z_2^2 + \ldots + Z_\nu^2$$

are o repartiţie χ^2 cu ν grade de libertate.

Se observă că singurul parametru al repartiţiei χ^2 este numărul gradelor de libertate. În general, prin **numărul gradelor de libertate** asociate unei entităţi se înţelege numărul variabilelor independente a căror variaţie nu suferă nici o restricţie şi care definesc mărimea luată în considerare.

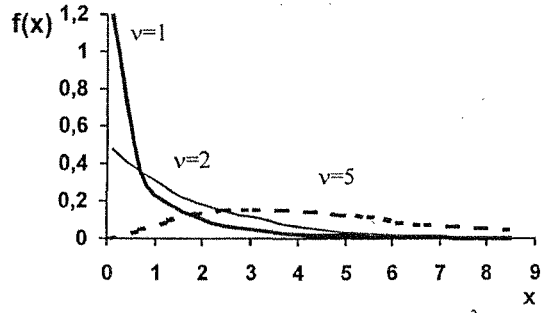


Figura 5.11. *Densităţi de probabilitate χ^2 pentru grade de libertate diferite*

Trebuie să observăm că notația variabilei aleatorii coincide cu denumirea repartiției, mod de notare utilizat și pentru alte repartiții.

În figura 5.11 sunt prezentate câteva densități de probabilitate a repartițiilor χ^2. Graficul este o curbă descrescătoare pentru $\nu = 1$ sau $\nu = 2$, iar pentru $\nu > 2$ are un punct maxim în punctul de abscisă $x = \nu - 2$.

Principalii indicatori statistici sunt:

$$M(\chi^2) = \nu,$$
$$D^2(\chi^2) = 2\nu.$$

Notăm quantila de ordin α a variabilei aleatorii χ^2 prin $\chi^2_{\alpha,\nu}$.

Deoarece nu există ramura negativă a repartiției, se obișnuiește să se noteze quantilele într-un mod inversat, adică

$$P(\chi^2 > \chi^2_{\alpha,\nu}) = \alpha.$$

Observație. După notațiile utilizate la celelalte repartiții, ar trebui să se folosească notația cu $1 - \alpha$ în loc de α.

Valorile quantilelor sunt tabelate pentru diferite valori ale lui α și diferite grade de libertate. În cazul în care $\nu > 30$, atunci se poate utiliza formula aproximativă

$$P(\chi^2 > \chi^2_{\alpha,\nu}) \approx 1 - F\left(\frac{1}{2}(z_\alpha + \sqrt{2\nu - 1})^2\right)$$

bazată pe funcția de repartiție normală normată.

5.7.4. Repartiția Student (t)

La fel ca repartiția precedentă, repartiția Student este utilizată în testarea ipotezelor statistice și a apărut datorită necesității de a compensa lipsa de informație privind atributele unei populații statistice prin acceptarea unor probabilități mai mari de eroare (afirmația va fi mai bine înțeleasă la prezentarea testelor statistice).

Repartiția Student depinde de un singur parametru, numit, ca și la repartiția χ^2, numărul gradelor de libertate și notat cu ν. Repartiția este denumită și repartiție t și sunt utilizate notațiile $S(\nu)$ sau t_ν.

Deoarece expresiile matematice ale funcțiilor de repartiție și densitate de probabilitate sunt relativ complicate și nu oferă nici un suport intuitiv suplimentar (implică o funcție neelementară, Γ), în figura 5.12 se prezintă, comparativ, graficele densităților de probabilitate normală standard și Student. Este luată în considerare repartiția t cu $\nu = 2$ grade de libertate.

Repartiția este simetrică față de 0 și se observă că, spre deosebire de repartiția normală, valorile extreme au probabilități mai mari de realizare (ariile de sub grafic sunt mai mari la repartiția t).

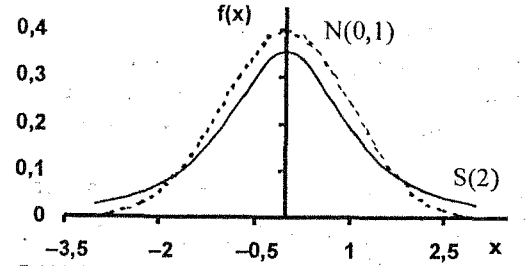


Figura 5.12. *Comparație între graficele densităților de probabilitate: N(0,1) – linia punctată; S(2) – linia continuă*

Indicatorii statistici principali ai unei v.a. $t \sim t_\nu$ sunt:

$$M(t) = 0;$$

$$D^2(t) = \frac{\nu}{\nu - 2}, \text{ pentru } \nu > 2.$$

Dacă $\nu > 30$, atunci $D^2(t)$ devine aproximativ egal cu 1, iar densitatea de probabilitate se apropie foarte mult de aceea a unei v.a. normale reduse. Acest fapt permite ca, pentru aplicațiile practice, repartiția Student cu mai mult de 30 de grade de libertate să fie aproximată cu o repartiție N(0;1).

Quantilele distribuției se notează cu $t_{\alpha;\nu}$, adică valorile pentru care

$$P(t < t_{\alpha;\nu}) = \alpha.$$

Aceste valori sunt tabelate pentru diferite valori ale lui α și ν. Pentru acele valori α și ν care nu figurează în tabel se utilizează aproximarea normală standard.

În aplicații este necesară uneori determinarea acelor valori ale distribuției t, care limitează arii extreme de valoare totală egală cu α. Situația este explicată în figura 5.13.

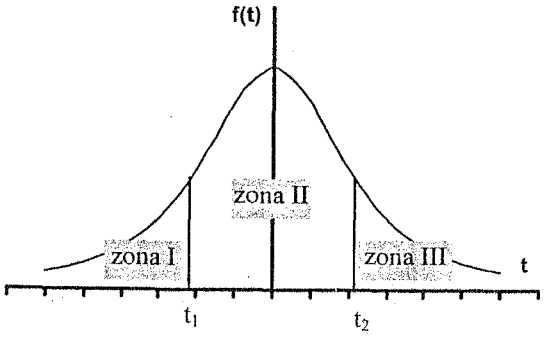


Figura 5.13. *Quantile care delimitează zone de interes*

a) Dacă aria zonei I este egală cu α, atunci $t_1 = t_{\alpha;\nu}$.
b) Dacă aria zonei III este egală cu α, atunci $t_2 = t_{1-\alpha;\nu}$.
c) Dacă aria zonei I împreună cu aria zonei III dau valoare egală cu α, fiecare fiind $\alpha/2$, atunci $t_1 = t_{\alpha/2;\nu}$ şi $t_2 = t_{1-\alpha/2;\nu}$. Situaţia coincide cu aria zonei II egală cu $1 - \alpha$.

Unele tabele statistice raportează şi asemenea valori, în special de tipul c).

Observaţie. Situaţiile reprezentate în figura 5.13 sunt aplicabile şi altor distribuţii, în special repartiţiei normale standard.

Dintre rezultatele importante care privesc repartiţia t menţionăm următoarea situaţie: dacă X este repartizată normal standard, X~N(0;1), iar Y este repartizată χ^2 cu ν grade de libertate, $Y \sim \chi^2_\nu$, atunci v.a. Z obţinută prin

$$Z = \frac{X}{\sqrt{\frac{Y}{\nu}}} = \frac{X \cdot \sqrt{\nu}}{\sqrt{Y}}$$

este repartizată Student cu ν grade de libertate.

5.7.5. Repartiţia Fisher-Snedecor (F)

Repartiţia F este introdusă ca o repartiţie utilă în testarea ipotezelor statistice privind compararea a două dispersii. Repartiţia F depinde de doi parametri, ν_1 şi ν_2, ambii având semnificaţia unor grade de libertate. Notaţia uzuală este $F(\nu_1,\nu_2)$ sau $F_{\nu1;\nu2}$.

Dacă repartiţia se referă la o variabilă aleatorie care are valoarea câtului altor două v.a., atunci ν_1 reprezintă numărul gradelor de libertate pentru numărător şi ν_2 – numărul gradelor de libertate pentru numitor.

Forma repartiţiei depinde de valorile parametrilor; o imagine orientativă este dată în figura 5.14.

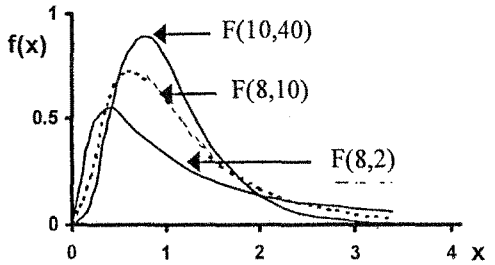


Figura 5.14. *Graficul densităţii de probabilitate F*

Pentru o v.a. F repartizată $F(\nu_1,\nu_2)$ se obţine:

$$M(F) = \frac{\nu_2}{\nu_2 - 2}, \text{ dacă } \nu_2 > 2;$$

$$D^2(F) = \frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 4)(\nu_2 - 2)^2}, \text{ pentru } \nu_2 > 4.$$

Quantilele repartiţiei F, notate $F_{\alpha;\nu1;\nu2}$, sunt tabelate pentru diferite valori uzuale ale lui α şi diferite grade de libertate. Repartiţia F nu este simetrică, între quantile existând relaţia

$$F_{1-\alpha;\nu_1;\nu_2} = \frac{1}{F_{\alpha;\nu_1;\nu_2}}.$$

Dintre rezultatele importante legate de repartiţia F menţionăm teorema următoare, luată uneori ca definiţie pentru o v.a. repartizată Fisher-Snedecor: dacă Y_1 şi Y_2 sunt două variabile aleatorii independente repartizate χ^2, cu ν_1 şi, respectiv, ν_2 grade de libertate, atunci v.a. X definită prin

$$X = \frac{\frac{Y_1}{\nu_1}}{\frac{Y_2}{\nu_2}} = \frac{Y_1 \cdot \nu_2}{Y_2 \cdot \nu_1}$$

are o repartiţie F cu ν_1 şi ν_2 grade de libertate.

5.7.6. Repartiţia binomială

Repartiţiile prezentate până acum în acest capitol sunt repartiţii definite pentru variabile aleatorii continue. Repartiţia binomială, denumită şi repartiţia lui Bernoulli, este definită pentru o variabilă discretă. Pentru ca o variabilă empirică să poată fi modelată printr-o v.a. repartizată binomial, este necesar să fie îndeplinite anumite condiţii.

Modelul binomial are ca punct de plecare schema bilei revenite:

Într-o urnă se află bile albe şi bile negre, în total N *bile. Probabilitatea de a extrage o bilă neagră este* p *(= numărul de bile negre raportat la numărul total de bile). După ce se extrage o bilă, aceasta se repune în urnă. Se cere probabilitatea ca în urma a* n *extrageri să se obţină* k *bile negre (şi, evident,* n – k *bile albe).*

Se numeşte **încercare** obţinerea unui rezultat individual şi **experiment** o suită de încercări.

Din analiza acestui model se pot reține condițiile de aplicare a modelului binomial:

- Variabila trebuie să fie dihotomică, adică să aibă doar două valori posibile (alb/negru, bărbat/femeie, încadrat/şomer etc.). Cele două valori sunt notate *succes/insucces.*
- Probabilitățile asociate valorilor *succes/insucces* trebuie să fie aceleași pentru tot experimentul. Notăm probabilitatea evenimentului de a obține valoarea succes cu p, adică P(succes) = p. Rezultă că P(insucces) = q = 1 – p (pentru că variabila este dihotomică).
- Încercările trebuie să fie independente între ele (obținerea unui rezultat anumit la o încercare nu influențează celelalte încercări).
- Încercările trebuie să fie identice prin aceea că procedura de obținere este aceeași pe toată durata experimentului. Este mai mult o recomandare practică urmărind păstrarea constanței probabilităților în experiment.

Variabila aleatorie X, considerată în modelul binomial, este numărul de succese apărute într-un experiment. Valorile posibile ale lui X, într-un experiment cu n încercări, sunt 0, 1, 2, ..., n.

Se poate arăta că, în condițiile date, v.a. X ia aceste valori cu probabilitățile:

$$P(X=r)=C_n^r p^r q^{n-r},\ r=1,2,\ldots,n,$$

unde p este probabilitatea unui succes, q=1 – p este probabilitatea unui insucces, iar n este numărul de încercări din experiment. Prin C_n^r se notează numărul de combinări de n obiecte luate câte r, mai exact:

$$C_n^r=\frac{n!}{r!(n-r)!}=\frac{1\cdot 2\cdot 3\cdot\ldots\cdot n}{1\cdot 2\cdot\ldots\cdot r\cdot 1\cdot 2\cdot\ldots\cdot(n-r)}.$$

Probabilitățile se mai notează prin p_r sau P(n;r), pentru a evidenția numărul de încercări.

Funcția de probabilitate a v.a. X este atunci imediată:

$$X:\begin{pmatrix}0 & 1 & 2 & \ldots & n\\ C_n^0 p^0 q^n & C_n^1 p^1 q^{n-1} & C_n^2 p^2 q^{n-2} & \ldots & C_n^n p^n q^0\end{pmatrix}.$$

Repartiția binomială se notează cu Bi(n;p), întrucât n şi p sunt parametrii definitorii ai funcției. Termenul de binomial este introdus pentru că probabilitățile P(n;r) se obțin ca termeni ai dezvoltării binomiale $(p+q)^n$.

Valorile tipice ale repartiției binomiale sunt:

$$M(X)=np,$$

$$D^2(X)=npq=np(1-p).$$

Funcția de repartiție este o funcție în trepte, specifică unei distribuții discrete:

$$F(x)=\sum_{k=0}^{x}P(n;k)=\sum_{k=0}^{x}C_n^k p^k q^{n-k},\ x\in\Re,$$

unde suma trebuie înțeleasă ca extinzându-se atâta timp cât k < x.

Atât valorile probabilităților P(n;r), cât şi valorile funcției de repartiție sunt tabelate şi se pot utiliza pentru calculul unor probabilități de interes.

Forma distribuției depinde de raportul dintre p şi q: distribuția este simetrică pentru p = q şi asimetria se accentuează o dată cu disproporția dintre p şi q. În figura 5.15 se poate observa simetria distribuției Bi(8;0,5), adică repartiția caracterizată de p = q = 0,5, precum şi asimetria pronunțată a distribuției Bi(5;0,8) în care p = 0,8, iar q = 0,2.

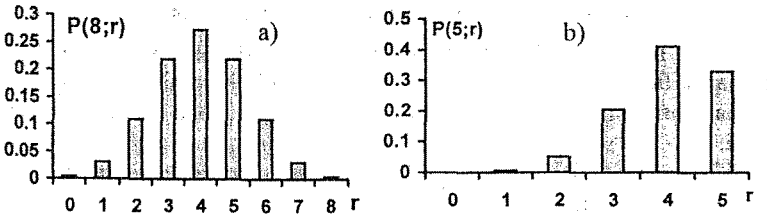


Figura 5.15. *a) Repartiția Bi(8;0,5); b) repartiția Bi(5;0,8)*

Pentru valori mari ale lui n, repartiția binomială poate fi aproximată prin repartiția normală cu N(np;npq). Condițiile sunt diferite, în funcție de autori; de exemplu:

- $np(1-p)>9$;
- $n\geq 30$şi una dintre condițiile $np\geq 5$ sau $nq\geq 5$;
- $np\geq 5$ şi $nq\geq 5$.

De remarcat că oricare dintre seturile de condiții necesită valori ale lui n suficient de mari (în general, cel puțin 30). Aproximarea serveşte la calculul probabilităților şi, deoarece o repartiție discretă este ajustată printr-o repartiție continuă, se utilizează, mai ales pentru valori mai mici ale lui n, corecțiile de continuitate:

$$P(X=r)=P\left(\frac{(r-np)-\frac{1}{2}}{\sqrt{np(1-p)}}\leq Z\leq\frac{(r-np)+\frac{1}{2}}{\sqrt{np(1-p)}}\right),$$

adică se aproximează probabilitatea ca v.a. X, repartizată Bi(n,p), să ia valoarea r prin probabilitatea ca v.a. Z, repartizată N(np, npq), să ia valori într-un interval de lungime egal cu unitatea, centrat în r. Utilizarea funcției de repartiție normală standard produce formula:

$$P(X=r)=F(z_2)-F(z_1),\text{ unde}$$

$$z_1=\frac{r+\frac{1}{2}-np}{\sqrt{np(1-p)}},\quad z_2=\frac{r+\frac{1}{2}-np}{\sqrt{np(1-p)}}.$$

Utilizarea practică a acestei formule presupune deci calculul prealabil al valorilor z transformate şi folosirea tabelului repartiţiei normale reduse.

De exemplu, pentru Bi(25, 1/2), din tabelul repartiţiei binomiale se citeşte $P(X = 9) = 0{,}0609$. Utilizând aproximarea normală, se obţine

$$P(X = 9) = P(-1{,}6 \le Z \le -1{,}2) = F(-1{,}2) - F(-1{,}6) = 0{,}060,$$

care constituie o bună aproximare a valorii exacte.

5.7.7. Aplicaţie la repartizarea observaţiilor în clase

S-a văzut că forma repartiţiei binomiale cu $p = q = 1/2$ este simetrică şi că, pentru un număr suficient de mare de observaţii, repartiţia binomială este aproximată de repartiţia normală. Aceste constatări pot servi în două situaţii importante în practică:

- verificarea apropierii măsurătorilor efectuate de o repartiţie normală, mai ales în cazul în care datele sunt grupate într-un număr de clase;
- gradarea unei scale continue atunci când dispunem de un număr suficient de mare de observaţii pentru o variabilă care se presupune că este continuă.

Într-o repartiţie binomială Bi(n,1/2), probabilităţile sunt proporţionale cu coeficienţii binomiali, deci cu numerele conţinute în **triunghiul lui Pascal**:

o clasă: 1
2 clase: 1 1
3 clase: 1 2 1
4 clase: 1 3 3 1
5 clase: 1 4 6 4 1
...

Construcţia triunghiului este recurentă, condusă de două reguli:

- prima linie are un singur număr: 1;
- o linie ulterioară începe şi se sfârşeşte cu 1, numerele intermediare sunt suma celor două numere situate pe linia anterioară deasupra şi la stânga:

a b
c = a + b

Să presupunem că dispunem într-o cercetare de 140 de observaţii şi am dori să ştim care ar fi efectivele teoretice pentru 5 clase de grupare, în cazul în care se acceptă o distribuţie normală pentru caracteristica studiată. Notând cu a, b, c, d şi e efectivele căutate, trebuie să aibă loc proporţionalitatea cu numerele situate pe linia a cincea a triunghiului lui Pascal:

$$\frac{a}{1} = \frac{b}{4} = \frac{c}{6} = \frac{d}{4} = \frac{e}{1} = \frac{a+b+c+d+e}{1+4+6+4+1} = \frac{140}{16} = 8{,}75$$

în care egalitatea primelor cinci rapoarte reflectă proporţionalitatea cerută.

Se aplică apoi regula de la şirurile de rapoarte egale prin completarea cu un raport egal cu suma numărătorilor pe suma numitorilor. Cum sunt 140 de observaţii

care se distribuie în 5 clase, rezultă că valoarea fiecărui raport este de 8,75. Egalând, pe rând, fiecare raport cu această valoare, rezultă:

$$a = e = 1 \times 8{,}75 = 8{,}75;\ b = d = 4 \times 8{,}75 = 35;\ c = 6 \times 8{,}75 = 52{,}5.$$

Rotunjind valorile obţinute, avem $a = 9$, $b = 35$, $c = 52$, $d = 35$, $e = 9$, care reprezintă efectivele teoretice ale celor 5 clase. Aceste efective pot fi utilizate în teste de comparare pentru a decide dacă repartiţia caracteristicii studiate se apropie de o repartiţie normală.

Calculele anterioare pot fi utilizate şi în scopul gradării unei scale continue: se decide (prin convenţie) că intervalele claselor sunt egale între ele.

Capitolul 6

INFERENȚA STATISTICĂ. ESTIMAȚII

Prin **inferență** înțelegem, potrivit definiției uzuale, *operația la nivelul intelectului prin care se trece de la un enunț la altul, acceptarea adevărului ultimului enunț bazându-se pe legătura logică menținută cu enunțurile anterioare, acceptate ca adevărate*. Prin inferență statistică se înțelege, în sensul precizat anterior, obținerea de concluzii bazate pe o evidență statistică, adică pe informații derivate dintr-un eșantion. Concluziile vizează caracteristicile populației din care provine eșantionul. Analizând definiția, se observă că inferența statistică reprezintă un întreg proces în care are loc o inducție, prin care cercetătorul trece de la cunoscut la necunoscut, de la particular la general.

Inferența statistică operează cu concepte proprii, care sunt prezentate în continuare.

6.1. Populație statistică

Prin **populație statistică** sau, simplu, **populație** vom înțelege mulțimea de entități sau activități de interes, care satisfac un criteriu prestabilit, propriu unei cercetări. O definiție precisă poate fi dată doar în legătură cu o cercetare fixată: dacă se studiază motivația pentru învățătură a studenților de la o anumită universitate, atunci mulțimea studenților acelei universități constituie populația statistică; dacă se studiază motivația pentru învățătură a studenților dintr-o anumită zonă geografică (regiune, țară etc.), atunci mulțimea studenților din acea zonă constituie populația statistică etc.

Observație. Dacă este investigată întreaga populație, atunci rezultatele obținute constituie finalul prelucrării și nu sunt necesare (și nici posibile) prelucrările introduse în această secțiune.

În statistică, termenul de populație are un sens aparte față de cel strict demografic (de ansamblu al locuitorilor unei zone geografice), sens dat de specificul acestei discipline, care operează cu mulțimi de date aparținând unor domenii foarte diverse. Ca atare, populația nu înseamnă totalitatea locuitorilor globului pământesc sau ai unei țări etc.

În continuare, termenul de populație va fi utilizat în sensul de populație statistică. Stabilirea criteriului potrivit căruia un element este inclus în populație

trebuie să fie foarte riguroasă, pentru a nu da naștere la confuzii. În loc de entități, se preferă în unele definiții ale populației termenii de indivizi sau de unități, cu precizarea că nu este vorba întotdeauna de indivizi umani, ci de elemente ale populației statistice. Făcând precizările de mai sus, putem da câteva exemple în acest sens:

- grupuri de persoane;
- grupuri de instituții;
- grupuri de animale;
- momente din istorie;
- grupuri de obiecte;
- grupuri de itemi.

Există o mare diversitate a modalităților în care putem să ilustrăm noțiunea de populație statistică. Criteriile de la care se pleacă în definirea individului de bază sau a unității formează un mare evantai, pornind de la cele mai banale până la cele mai „ciudate". Astfel, exemplificând în domeniul uman, într-o anumită populație pot fi cuprinse:

- toate persoanele de naționalitate română;
- toate persoanele care sunt angajate într-o întreprindere;
- toate persoanele care folosesc un anumit tip de șampon;
- toate persoanele dintr-un anumit județ care posedă un câine de rasă;
- toți inginerii care sunt directori de întreprindere.

Practic, nu există limite ale definirii caracteristicilor unei populații. Din punctul de vedere al statisticii, calitățile pe care trebuie să le aibă o populație trebuie bine precizate și, în majoritatea cazurilor, clar cuantificate. Mărimea unei populații, dată prin numărul elementelor componente, poate varia de la un individ sau o unitate (mulțimea papilor în viață) până la numere extrem de mari. În unele cazuri, când numărul unităților este foarte mare și se găsește într-o continuă creștere, cu o rată imprevizibilă, se vorbește de populații infinite. Același individ sau aceeași unitate poate să se regăsească în mai multe populații, în cazul în care satisface și criteriile altor sisteme de clasificare. Astfel, o persoană poate face parte din populația profesorilor de biologie din Moldova, din populația credincioșilor ortodocși din aceeași zonă, din populația posesorilor de computere, din populația actorilor amatori etc.

În cazul în care se calculează un indicator numeric din date provenind de la toți membrii populației, mărimea rezultată se numește **parametru** și are o valoare fixă, nesupusă oscilației (cel puțin până la modificarea mărimii populației). Este doar unul dintre sensurile noțiunii de parametru. Dacă avem posibilitatea să măsurăm în întregime înălțimea populației de copii de sex masculin, în vârstă de 12 ani, din România, putem spune cu exactitate cât este media de înălțime a populației statistice studiate doar în momentul efectuării studiului. Un asemenea studiu este practic imposibil, deoarece ar presupune ca, la o aceeași dată și la aceeași oră, să fie măsurată înălțimea oricărui copil din populația menționată. Din acest motiv, o cercetare se efectuează pe submulțimi ale populației statistice.

6.2. Eşantion

Prin **eşantion** (sau **selecţie**) vom înţelege o submulţime a populaţiei statistice avute în vedere. O cercetare statistică are întotdeauna ca obiect un eşantion finit. De obicei, termenul ca atare presupune că eşantionul este şi reprezentativ, adică informaţia aferentă lui este în concordanţă cu informaţia caracteristică populaţiei statistice. Reprezentativitatea unui eşantion depinde de modul de selectare a elementelor constitutive şi nu există, practic, un instrument cu ajutorul căruia să se spună că un eşantion este sau nu reprezentativ în absenţa informaţiilor privind populaţia (a se vedea observaţia anterioară). Ceea ce se poate face, în mod raţional, este să se aleagă acele metode de constituire a unui eşantion care au şanse mai mari să producă eşantioane reprezentative.

Un eşantion serveşte ca bază în măsurătorile utilizate pentru a trage concluzii în privinţa întregii populaţii din care a fost extras. Măsurătorile efectuate asupra eşantioanelor pot oferi doar estimaţia valorii parametrilor, estimaţie ce poate fi destul de apropiată de valorile adevărate, valorile populaţiei, în cazul în care se parcurg paşii necesari pentru constituirea unei metodologii riguroase de constituire a eşantionului. Un eşantion este compus din „unităţi de sondaj". Mărimea unui eşantion desemnează numărul de unităţi componente. În anchetele psihosociologice, unităţile sunt, în general, persoane umane. Pentru alte genuri de investigaţii, unităţile pot fi o situaţie, un document, o unitate de înregistrare video sau audiovizuală, un articol de ziar etc.

Operaţia de formare a unui eşantion se numeşte **sondaj** sau **eşantionaj**. Tehnicile de eşantionaj sunt identice, oricare ar fi natura unităţilor eşantionului. Sondajele care au şanse mai mari de a produce eşantioane reprezentative sunt cele bazate pe proceduri de selecţie aleatorii.

Exemplul 6.1. Într-un studiu asupra caracteristicilor memoriei copiilor din clasele primare, se va putea observa acest proces psihic cognitiv în mai multe clase primare din România pentru a evidenţia modul în care se efectuează memorarea sau ce anume caracterizează geneza şi evoluţia acesteia (realizarea studiului cuprinde, cu siguranţă, detalii de operaţionalizare conceptuală care nu prezintă interes pentru scopul exemplificării noastre). Se va studia, în acest caz, un eşantion din populaţia elevilor din România, având vârsta stabilită de către cercetători înaintea efectuării studiului. Dacă acelaşi studiu se va efectua asupra unui eşantion sau asupra mai multor eşantioane de elevi de aceeaşi vârstă, din şcolile unei ţări din Africa şi ale unei ţări din America Latină, vor exista eşantioane provenind din trei populaţii diferite.

Întrebarea care se pune în acest caz este următoarea: dacă se observă diferenţe fundamentale de memorare între două eşantioane de elevi, provenind din ţări diferite, este posibil ca diferenţele respective să existe şi la nivelul populaţiilor luate în întregime? Răspunsul la această întrebare este foarte important, deoarece, dacă se pot confirma diferenţele, se trage concluzia existenţei diferenţelor între

populaţii. Dacă nu se poate realiza acest lucru, nu se pot trage astfel de concluzii (în ciuda diferenţelor observate între eşantioane) şi se poate conchide că geneza memoriei şi a diferitelor sale forme este universală, că ea prezintă aceleaşi faze şi etape la toţi copiii.

În alte conjuncturi şi în alte situaţii, miza nu poate fi la fel de importantă. Apare în acest caz obligaţia de a arăta că ceea ce s-a obţinut sau s-a observat nu este pur fortuit şi ocazional, propriu doar eşantionului studiat, ci că avem de-a face cu lucruri semnificative, suficient de generale, valabile pentru populaţie, adică situate dincolo de nivelul eşantionului. Faptul observat are un caracter generalizabil, aplicabil nu la toţi indivizii, nici la toate eşantioanele de indivizi, ci la cea mai mare parte a acestora.

Demersul constând în a generaliza pentru un ansamblu mai vast (populaţia) ceea ce se observă pe un ansamblu mult mai restrâns (eşantion) este cel mai uzual în psihologie, pentru că, în general, parametrii descriind populaţia nu sunt cunoscuţi.

Mai rar, când se cunoaşte populaţia, se pune problema de a şti dacă un eşantion observat este conform, tipic acestei populaţii, reprezentativ, sau dacă este atipic şi, probabil, rezultat din altă populaţie decât cea vizată de studiul respectiv. O astfel de situaţie de cercetare se întâlneşte când se testează un model (elaborat de un cercetător) prezicând răspunsurile aşteptate din partea subiecţilor. Este vorba despre a şti dacă răspunsurile efective obţinute pe un eşantion de subiecţi (plasaţi în condiţii definite prin model) sunt conforme cu predicţiile acestuia.

În ambele cazuri discutate, în care nu cunoaştem sau cunoaştem parametrii populaţiei, pentru a se pronunţa în problematica amintită, cercetătorul apelează la utilizarea tehnicilor statisticii inferenţiale.

6.3. Clasificarea populaţiei şi a eşantioanelor

Dacă urmărim criteriul de formare, populaţia şi eşantioanele pot prezenta o mare diversitate. Sunt discutate în continuare doar clasificările cu care se operează în domeniul psihologiei şi pedagogiei. În cele mai multe cazuri ale cercetărilor ştiinţifice din aceste domenii nu se operează cu un singur eşantion, ci cu mai multe. Cele mai frecvente referiri se fac după criteriul mărimii şi după cel al dependenţei/independenţei elementelor.

Fixarea mărimii eşantionului ce urmează să fie constituit este o problemă importantă. Numărul persoanelor selectate, volumul eşantionului, trebuie să satisfacă cerinţa de a fi suficient de mare pentru a putea efectua estimarea mai exactă a populaţiei şi, în acelaşi timp, să fie suficient de mic pentru a nu depăşi costurile fixate pentru cercetare. Persoanele selectate trebuie, evident, să fie capabile să ofere informaţii fiabile, ţinând cont de caracteristicile de interes. De exemplu, în cazul sondajului de opinii asupra intenţiilor de vot, este important ca persoanele să fie chestionate într-un interval temporal foarte scurt, pentru că se ştie că opţiunile electorale sunt foarte fluctuante şi că anticiparea rezultatului votului constituie interesul

major al operației. Se limitează numărul de persoane chestionate, datorită costului sondajului și numărului de anchetatori specializați disponibili într-un timp scurt.

Din punctul de vedere al mărimii, se disting eșantioane mici și eșantioane mari. Prin convenție, dar și din considerente teoretice, este acceptată în majoritatea cazurilor cifra 30 drept graniță între cele două tipuri. Dar, ca în majoritatea convențiilor, se produc și în acest caz încălcări ale regulii. În domeniul nostru de interes, se constată uneori „dilatări" ușoare ale eșantioanelor mici, fiind uneori limitate la 50. În orice caz, a trata un eșantion mic ca pe unul mare poate produce erori importante, în timp ce tratarea unui eșantion mare ca pe unul mic nu este foarte grav. Clasificarea este importantă, deoarece multe proceduri de analiză sunt adecvate unui tip de eșantion.

Dacă în cercetare sunt implicate mai multe eșantioane, acestea se clasifică în independente și dependente (după modul de selectare a elementelor). Două eșantioane se numesc **independente** în cazul în care selectarea unui element într-un eșantion nu depinde de elementele celuilalt eșantion. În psihologie, cazul tipic de eșantioane independente este acela în care indivizii care le compun nu sunt fizic aceiași și nu sunt selectați prin urmărirea unui criteriu suplimentar în fiecare dintre ele. Ilustrarea cea mai clară a unor astfel de eșantioane sunt cele diferențiate prin sex. Aceste tipuri de eșantioane se folosesc atunci când cercetătorul vrea să evite ca efectele adiționale ale manipulării modalităților variabilelor să se producă, atunci când există o incompatibilitate absolută în succesiunea de manipulare a acestor modalități, când se dorește evitarea efectelor oboselii, a scăderii motivației și a atenției din partea indivizilor care le compun și în cazul în care există o mare solicitare în cadrul unor cercetări.

Eșantioanele **perechi** sau **dependente** sunt formate din aceiași indivizi sau din grupuri diferite de indivizi, care au fost anterior egalizate în privința unor anumite caracteristici (vârstă, sex, nivel de școlarizare, arie geografică de proveniență etc.) și au fost pregătite cu grijă pentru competiție. Selectarea unui individ într-un eșantion conduce automat la selectarea unui individ cu caracteristici similare în celălalt eșantion. Utilizarea eșantioanelor dependente în psihologie și pedagogie se înscrie în trei categorii de studii experimentale:

- Cercetări în care există necesitatea de a controla cât mai riguros un factor de variație. În cazul cercetărilor longitudinale, același eșantion este folosit mai mulți ani sau chiar mai multe decenii. În investigațiile științifice de medicină psihosomatică, cercetătorii utilizează aceiași indivizi la un interval de 20 sau 30 de ani, după ce anterior au constatat la aceștia prezența unor factori de ordin psihologic care predispun la îmbolnăvirea cardiacă sau de altă natură. Firește, în asemenea studii apar probleme specifice, cum ar fi reducerea volumului eșantionului. Identitatea nu se caută pe caracteristicile comune eșantioanelor, ci pe identitatea proprie indivizilor. Deci este vorba de aceiași indivizi pe care se prelevează mai multe măsurători.
  În cercetările psihosociologice este frecvent folosit, pentru acest tip de cercetare, termenul de *panel*, mai ales pentru anchetele diacronice. Când interesează un proces evolutiv, este necesar să se interogheze, la date succesive, același eșantion, denumit *panel*. Metoda a fost descrisă de Paul Lazarsfeld în anul 1938. Ea este utilizată mai ales pentru studiul evoluției comportamentelor electorale. Institutele de sondaje au *panel*-uri permanente pe subpopulații particulare: medici, avocați, auditori de radio, televiziune, grupări religioase și politice. *Panel*-urile sunt de asemenea utilizate atunci când un cercetător are nevoie să verifice experiența sau competența indivizilor. Pentru această tehnică prezintă interes reperarea indivizilor care se transformă și a celor care nu se transformă, înțelegerea nivelului acestei stabilități sau evoluții. Metoda prezintă unele inconveniente, deoarece persoanele care acceptă să facă parte dintr-un panel au caracteristici particulare (disponibilitate, sociabilitate) și este dificil să se diferențieze în ceea ce influențează comportamentele, opiniile sau atitudinile. Atitudinile se pot schimba prin faptul... că ești membru al unui *panel* (de exemplu, tendința de a da răspunsuri conformiste, de a rămâne coerent în timp etc.). De asemenea, există un risc de denaturare a *panel*-ului, fenomen datorat evoluției persoanelor (îmbătrânire, schimbarea profesiei, încheierea colaborării, schimbarea locuinței).
- Cercetări și studii în care diferitele etape de măsurare sunt legate între ele, în cazul în care un cercetător își propune să aprecieze efectele cumulative ale manipulării variabilelor. Exemple de studii de acest gen sunt cele efectuate în privința efectului învățării asupra memoriei.
- Cercetări în care se dorește controlarea influenței factorilor comuni eșantioanelor. În acest tip de cercetare, nu interesează caracteristicile comune ale indivizilor, ci acelea care sunt comune grupurilor.

6.4. Eșantionarea (sondajul)

Maniera în care se procedează pentru constituirea unui eșantion depinde de urmărirea unor obiective. Ca atare, este important să știm:

- dacă avem informații asupra incidenței caracteristicilor definite în prealabil. De exemplu: bărbații și femeile au aceleași opțiuni în privința profesiei care urmează a fi practicată? Există o diferențiere a acestor opțiuni în funcție de zona geografică, de mediul social de proveniență, de vârsta la care se face opțiunea? Cel mai bine informat și cu experiență într-un anumit domeniu are mari avantaje în efectuarea eșantionării din domeniul respectiv;
- dacă există informații asupra populației, altfel spus, dacă se cunosc parametrii acesteia. Desigur, ne putem întreba în mod firesc dacă mai avem nevoie de eșantioane, de vreme ce cunoaștem caracteristicile populației. Cunoașterea indicatorilor statistici ai unei populații nu înlătură nevoia constituirii de eșantioane, deoarece în unele cazuri avem nevoie să știm dacă aceste eșantioane sunt reprezentative pentru populația respectivă (în scopul efectuării unor alte cercetări).

Este important, în primul caz, ca grupele comparate să fie egalizate, cu excepția variabilelor definite (variabilele anchetei). În al doilea caz, este imperativ ca grupele reținute să fie constituite în așa fel încât ele să reprezinte populația de interes (populația părinte), deoarece se vor generaliza, prin inferență, la respectiva populație (populația părinte) constatările noi obținute pe baza datelor din eșantion.

Se impune ca un astfel de eșantion să fie reprezentativ. În caz contrar, se spune că eșantionul este deformat sau deplasat – *biais*[46]. Întotdeauna există diferențe între populație și eșantioane, dar acestea sunt minimizate prin tehnicile de eșantionare, care vizează obținerea, pe cât posibil, a unui eșantion reprezentativ.

Metodele de construcție a eșantioanelor sunt de două feluri: probabiliste și empirice. Primele metode se bazează pe selectarea aleatorie (cum ar fi tragerea la sorți) a elementelor constitutive. Constituirea eșantioanelor prin metode empirice se bazează pe cunoașterea structurii populației.

6.4.1. Sondaje aleatorii

În continuare sunt prezentate eșantionările probabiliste, bazate pe principiul alegerii aleatorii:

● **Sondaj probabilistic** sau **sondaj la întâmplare** (*random sampling*), obținut prin tragere la sorți elementară[47]. Acest tip de eșantion face apel la mecanismele de extragere aleatorie (prin tragere la sorți) a indivizilor care compun eșantionul. Aceasta presupune că cercetătorul dispune de date referitoare la ansamblul populației, adică de o **bază de sondaj**. Totuși, nu este cel mai adaptat studiilor în cazul în care unitățile de eșantionaj astfel selecționate nu sunt disponibile, mai ales când este vorba de persoane (este mai ușor de format un eșantion de persoane pentru a le măsura o dimensiune fizică decât un eșantion statistic pentru a măsura atitudinile). Tragerea la sorți elementară constă în a numerota toți indivizii care compun populația studiată, apoi de a extrage la întâmplare o cantitate definită din aceste numere; indivizii cărora li se atribuie numerele extrase alcătuiesc eșantionul.

Tragerea la sorți se poate efectua și cu ajutorul unui tabel de numere aleatorii (tipărit sau generat de un program specializat de calculator). Indiferent de procedura utilizată în extragere, este obligatoriu ca fiecare individ să aibă aceeași șansă de a fi ales.

Este modul de eșantionaj tipic probabilist, dar inconvenientul major constă în faptul că este foarte costisitor în timp și bani. A inventaria toți indivizii populației este o operație deosebit de dificilă, urmată de o alta la fel de dificilă, aceea de a contacta toți indivizii trași la sorți. Acest tip de sondaj se poate utiliza totuși în studiile de teren, în studiile efectuate pe copiii dintr-o creșă, în studiul efectuat într-o întreprindere sau, în general, în studiile efectuate pe mici colectivități.

46. Termenul *biais* indică o deformare sistematică, neîntâmplătoare, a seriei de date dintr-o cercetare. Cuvântul este preluat ca atare în multe studii datorită dificultății de a găsi cel mai potrivit termen românesc corespunzător.

47. Totuși, în majoritatea cazurilor, tragerea la sorți va fi precedată de o prestructurare empirică (deci nealeatorie), din rațiuni de cost și de fezabilitate.

● **Eșantionare prin tragere la sorți secvențială** sau **sondaj cu mai multe trepte** (*multiple stage sampling*). Este o tragere la sorți care se efectuează pe criterii succesive. În primul rând, se definește o populație pe un criteriu dat și se extrage un prim eșantion, urmând criteriul respectiv. Apoi se definește un al doilea criteriu, în interiorul acestui eșantion, și se procedează la o nouă tragere la sorți printre elementele inițial selecționate. Procedeul poate fi continuat pe mai multe niveluri.

Exemplu de tragere la sorți secvențială: studiul fenomenului de sugestibilitate prin reclamă. Se selecționează inițial un număr de localități (prima fază), apoi se extrag eșantioane de străzi din fiecare localitate selecționată (faza a doua) și, în final, de pe fiecare stradă se selecționează un număr de persoane (faza a treia).

● **Eșantionare grupată** (*grappes* sau *cluster sampling*). Un *grappe* este un ansamblu de unități având în general o legătură „naturală". De exemplu, militarii unui pluton, plutoanele aceleiași companii, companiile aceluiași batalion, elevii unei clase, clasele aceluiași așezământ școlar, muncitorii unui atelier, atelierele unei aceleiași uzine, o uzină din cadrul aceluiași grup de uzine având același profil, articolele unui ziar, ziarele unui aceluiași grup de presă. Unitățile componente ale unui *grappe* se aseamănă mai mult între ele decât cu alte unități. Baza de sondaj nu mai este atunci ansamblul unităților susceptibile de a fi selectate, ci *grappe*-urile. Eșantionarea obținută este o variantă a sondajului secvențial. În anchetele asupra școlii, este clasic să se procedeze mai întâi la tragerea la sorți a instituțiilor școlare, apoi a claselor, apoi a elevilor în interiorul nivelurilor școlare. Este vorba în acest caz de o selecție prin *grappes* cu trei etape sau trei trepte, pentru că sunt utilizate trei unități de sondaj legate. Prima treaptă se numește treapta primară, a doua treapta secundară, a treia terțiară ș.a.m.d. O eșantionare care s-a oprit la treapta a doua se va numi *eșantionare cu două trepte* (*two stage sampling*). Regulile de extragere pot fi identice sau diferite pentru diverse trepte. Avantajul acestui procedeu rezidă în faptul că un cercetător nu mai are nevoie să dispună de baza de sondaj privind populația și în faptul că se limitează dispersia unităților de contactat. Trebuie totuși ca *grappe*-urile să fie de mărime considerabilă și în număr suficient pentru a avea loc o tragere la sorți efectivă.

● **Eșantionare stratificată.** În acest caz se divizează populația în subpopulații denumite straturi, care dispun fiecare de caracteristici distincte. Procedura ce trebuie să fie urmată constă în a extrage un eșantion din interiorul fiecărui strat. Ansamblul acestor eșantioane, care pot fi de mărime diferită, constituie eșantionul de studiu.

Metoda presupune definirea în prealabil a subpopulațiilor sau straturilor, prin luarea în considerare a unuia sau a mai multe criterii de stratificare. Un strat definit corespunde unei modalități (cazul unui criteriu unic) sau unei intersecții de modalități (cazul mai multor criterii). Alegerea prin tragere la sorți a unităților se face apoi în interiorul fiecărui strat. Criteriile de stratificare a unităților se stabilesc în funcție de studiul efectuat. De exemplu, într-o anchetă asupra comportamentelor de

cumpărare a bunurilor de consum se iau drept criterii nivelul de venituri al cumpărătorului, sexul, mărimea aglomerației urbane etc. Într-o anchetă asupra funcționarilor se pot lua în considerare domeniul în care activează, sexul, vârsta, tipul de studii.

6.4.2. Sondaje nealeatorii

Metodele de eșantionare empirică sau metodele nealeatorii se bazează pe construirea, pornind de la informațiile despre structura populației, a unui eșantion a cărui structură este apropiată de cea a populației. Structura este definită prin distribuția procentajelor unei variabile categoriale.

Eșantionaj prin cote (*quote*). Este cel mai familiar tip de eșantionaj și constă în alegerea indivizilor care compun eșantionul în așa fel încât caracteristicile esențiale ale populației să se regăsească în aceleași proporții (cote) și în eșantion. În acest fel se construiește un „model redus al populației". Mărimea diferitelor subgrupuri constitutive ale eșantionului este proporțională cu cea a subgrupurilor echivalente din populație. Pentru acest tip de eșantionare se mai utilizează denumirea de „eșantionare proporțională". Dacă, de exemplu, în populație există 48% bărbați și 52% femei, într-o anchetă efectuată pe 1000 de persoane se vor chestiona 480 de bărbați (adică 48% din eșantion) și 520 de femei (adică 52% din eșantion). Cotele sunt fixate în prealabil și sunt comunicate persoanelor care se ocupă cu colectarea informațiilor științifice. Indivizii sunt în general greu de găsit, dacă ei aparțin unei categorii puțin frecvente din ansamblul populației. Se examinează indivizii la concurență cu aceste proporții și cu criterii fixate anterior. Cele mai importante criterii pentru adulți sunt: vârsta, sexul, profesia, nivelul de studii, nivelul de venituri, zona locuită. La copii se iau în considerare mai ales vârsta și echivalentul său (clasa), tipul de școală, profesiunea tatălui etc. Se pot colecta date până când se completează procentajele stabilite într-o grilă. În sondajele de opinie, în care metoda cotelor este foarte utilizată, variabilele avute în vedere mai frecvent sunt sexul, vârsta, mărimea comunității din care face parte subiectul, localizarea geografică, profesiunea capului de familie, profesiunea persoanei anchetate, ca și nivelul de studii. Fiabilitatea eșantionului se stabilește în funcție de justețea descrierii populației și de pertinența variabilelor alese pentru a determina cotele în legătură cu finalitatea studiului. În domeniile noi de studiu și în timpul bulversărilor sociale sau politice, este adesea dificil să se determine *a priori* variabilele pertinente.

Exemplul 6.2. Să presupunem că trebuie să constituim un eșantion de volum egal cu 300 de persoane în vederea unei examinări psihologice. La selectarea unităților trebuie să ținem seama de variabilele sex, vârstă și nivel de studii. Notăm vârsta cu V și luăm modalitățile: V_1 = 18-25 ani, V_2 = 26-40 ani, V_3 = 40-55 ani și V_4 = 56 ani și mai mult. Sexul se notează cu S și se iau modalitățile: S_1 = masculin și S_2 = feminin. Nivelul de studii îl vom nota cu G cu modalitățile: G_1 = studii generale, G_2 = studii medii și G_3 = studii superioare. Se întocmește apoi un tabel cu

24 de variante distincte ($4 \times 3 \times 2=24$) în care se trec cotele procentuale necesare (obținute din datele oficiale sau din cercetări anterioare). Fiecare casetă cuprinde cota procentuală așteptată și, în paranteză, efectivul corespunzător. Astfel, efectivul casetei $V_2S_1G_1$ se obține din operațiile $8{,}8\times300/100=26{,}4$ (aproximativ 26), reprezentând numărul de subiecți din eșantion, corespunzător cotei existente în întreaga populație.

În continuare se dau tabelele completate cu o situație fictivă de cote procentuale (tabelul din stânga) și efectivele corespunzătoare (tabelul din dreapta):

		V_1	V_2	V_3	V_4
G_1	S_1	8,4	8,8	8,6	6,3
	S_2	8,7	9,1	8,9	6,5
G_2	S_1	3,4	3,6	3,5	2,6
	S_2	3,5	3,6	3,5	2,6
G_3	S_1	1,1	1,1	1,1	0,8
	S_2	1,1	1,2	1,2	0,8

		V_1	V_2	V_3	V_4
G_1	S_1	25	26	26	19
	S_2	26	27	27	20
G_2	S_1	10	11	11	8
	S_2	11	11	11	8
G_3	S_1	3	3	3	2
	S_2	3	4	4	2

Aplicarea acestei metode de eșantionare constă în selectarea indivizilor până la atingerea efectivelor stabilite pentru fiecare casetă. Acest lucru se realizează prin mărirea ariei de investigație până când obținem valorile stabilite sau, dacă nu se pot atinge anumite valori, oprim selecția și se determină printr-un test statistic adecvat dacă valorile eșantionului obținut se distanțează semnificativ de cele așteptate.

Eșantionaj prin unități tip. Acest mod de sondare se obține prin divizarea populației în categorii cu ajutorul variabilelor considerate esențiale în problematica cercetării respective. Obiectivul este acela de construire a unui eșantion susceptibil de a furniza informații pertinente. De exemplu, o cercetare asupra inadaptării profesionale se va efectua în primul rând asupra persoanelor considerate ca inadaptate profesional. Este necesar în acest caz să se stabilească clar conținutul noțiunii de inadaptare profesională. Alte cercetări de acest gen sunt realizate, de exemplu, asupra infractorilor juvenili.

6.4.3. Erorile sistematice în sondaj

Dacă unitățile de eșantionaj sunt ființe umane, se fac în mod frecvent erori sistematice (*biais*-uri). În fond, nici o persoană nu este obligată să furnizeze informații cercetătorului; de aceea, în cvasitotalitatea cazurile se apelează la bunăvoința persoanelor respective. În unele cazuri, se poate apela la instituții care efectuează examene psihologice obligatorii, de exemplu laboratoare pentru examinarea psihologică a conducătorilor auto. Deformări ale informațiilor de eșantionaj există însă, oricare ar fi metodele folosite:

- în cazul unui eșantionaj statistic, o parte din indivizii care fac parte din eșantionul prevăzut nu răspund la chestionar sau la solicitarea cercetătorului;

- în cazul unui eşantionaj prin cote, se poate ajunge la o subreprezentare a unor anumite categorii de persoane.

Erorile sunt în mod evident mari în cazul în care se chestionează persoane prin telefon (cele care nu au telefon sunt ignorate), în locuri frecventate în manieră privilegiată de anumite categorii de persoane etc.

Pentru a controla aceste deformări informaţionale, se impun diferite reguli anchetatorilor: nu se chestionează prieteni, rude sau cunoştinţe, se schimbă locurile şi orele de anchetă, nu se chestionează mai multe persoane aparţinând aceluiaşi grup (membrii unei aceleiaşi familii, elevii aceloraşi clase) dacă nu se prevede aşa ceva. Institutele de sondaje exercită un control asupra anchetatorilor, deoarece ştiu că nivelul unei anchete depinde de respectarea regulilor de eşantionare, care au fost concepute în funcţie de finalitatea studiului. În ciuda acestor controale, ţinând cont de diversitatea surselor de eroare, se estimează că metoda cotelor dă în general erori în jur de 10%. Când este posibil să se aprecieze importanţa erorilor, se procedează la corecţii ale datelor obţinute. Corecţia nu este posibilă decât dacă există o bună cunoaştere a populaţiei investigate şi modele adecvate de redresare a datelor.

Redresarea eşantioanelor este o operaţie ce presupune o serie de riscuri. În chestionare se introduce adesea o întrebare (sau mai multe întrebări) care vizează sinceritatea răspunsurilor. De exemplu, când se investighează o intenţie de vot, se chestionează şi cu privire la votul precedent. Dacă există distorsiuni mari, se operează o redresare şi o ponderare.

Reamintim că valorile tipice ale distribuţiilor din populaţie reprezintă parametrii populaţiei, iar valorile calculate într-un eşantion sunt numite statistici. Datorită modului de obţinere, aceste statistici mai sunt numite *„de sondaj"*: *media* este, în general, *media populaţiei*, în timp ce *media de sondaj* este *media eşantionului*. Atunci când nu pot să apară confuzii, calificativul „de sondaj" este omis pentru a nu îngreuna discursul.

6.5. Acurateţe şi eroare în sondaj

Un eşantion reprezintă populaţia cu acurateţe dacă statistica eşantionului este egală cu parametrul populaţiei. De exemplu, dacă media de înălţime a populaţiei este de 1,63 m şi dispunem de un eşantion cu media de înălţime tot de 1,63 m, acesta reprezintă exact populaţia. Apar erori atunci când nu se întâmplă acest lucru. Erorile pot fi, după modul de apariţie, **sistematice** şi **întâmplătoare.**

Erorile sistematice apar, de cele mai multe ori, ca erori instrumentale, datorate deci instrumentelor utilizate, gradului de precizie a instrumentelor. O procedură de selecţie eronată poate fi o altă cauză de erori sistematice. Aceste erori pot fi înlăturate doar prin revizuirea procedeelor de sondaj sau/şi a instrumentelor de măsură.

Erorile întâmplătoare sau neregulate sunt cele care se produc datorită observatorului în procesul de măsurare (erori subiective) sau cele datorate unor influenţe

necontrolabile ce pot să apară în timpul măsurătorilor. Erorile întâmplătoare nu pot fi, în general, eliminate. Trebuie precizat că repartiţia normală a apărut tocmai ca urmare a studiului erorilor întâmplătoare efectuat de Gauss.

Chiar dacă se elimină toate erorile sistematice, chiar dacă se utilizează o procedură corectă de sondaj, întotdeauna este de aşteptat să existe diferenţe între parametrul populaţiei şi statistica eşantionului. De exemplu, la aruncarea unei monede corecte, şansa de a obţine în 8 aruncări o configuraţie 4 – 4 între feţele monedei este doar de 27% (cum se poate verifica în tabelul repartiţiei binomiale Bi(8,1/2)), deşi teoretic ar trebui ca probabilitatea să fie de 50%. Din aceeaşi distribuţie se poate vedea că sunt posibile şi configuraţii 8 – 0 sau 0 – 8, chiar în condiţii corecte de probă.

Apare astfel ideea importantă că *în eşantioane diferite statisticile calculate au valori diferite.* În acest fel, se poate vorbi despre o distribuţie a valorilor statisticii în mulţimea eşantioanelor cu un acelaşi volum. Este evident că, dacă se cunoaşte această distribuţie de valori, procesul de validare a cercetărilor se va desfăşura în condiţii mai bune.

În practică, din păcate, nu se dispune de o astfel de distribuţie. Cercetătorul doreşte să selecteze un eşantion şi, din prelucrarea acestuia, să obţină estimări ale parametrilor populaţiei. În plus, bazat tot pe informaţia disponibilă în eşantion, trebuie să decidă asupra acurateţei estimaţiilor obţinute.

6.6. Logica inferenţei statistice

Inferenţa statistică implică trei distribuţii asociate caracteristicii studiate:

a) distribuţia populaţiei;
b) distribuţia de sondaj;
c) distribuţia eşantionului.

Prin distribuţia populaţiei se înţelege distribuţia pe care o are caracteristica studiată (sau v.a. asociată ei) în populaţie. Această distribuţie nu este, în general, cunoscută. Interesul unei cercetări este tocmai acela de a studia o astfel de distribuţie.

Prin distribuţia eşantionului se înţelege distribuţia pe care o are caracteristica studiată în eşantionul disponibil în studiu. Este cunoscută complet, întrucât toate datele necesare sunt disponibile prin măsurare.

Prin distribuţia de sondaj se înţelege distribuţia pe care o are statistica în mulţimea tuturor eşantioanelor de volum dat. Această distribuţie nu este cunoscută şi conţine „mai multe" valori decât populaţia de bază (dacă populaţia de bază este infinită, sensul matematic pentru „mai multe" necesită o discuţie specială). Astfel, dacă populaţia de bază are N = 100 elemente, atunci există

$$C_{100}^{3} = \frac{100!}{3!\times 97!} = \frac{98\times 99\times 100}{1\times 2\times 3} = 161700$$

eşantioane de volum 3. Prin C_{100}^3 se notează numărul combinărilor de 100 de elemente luate câte 3. Se observă astfel că numărul celor 161700 de eşantioane posibile este mult mai mare decât numărul unităţilor din populaţia de bază (100). Numărul creşte foarte repede pe măsură ce populaţia este mai vastă şi volumul eşantionului se măreşte. Este însă remarcabil faptul că, din considerente teoretice, între distribuţia populaţiei şi distribuţia de sondaj există legături bine precizate sau, datorită unor teoreme de limită centrală, se cunoaşte forma acestei distribuţii atunci când volumul eşantionului creşte (teoretic, tinde spre infinit).

Inferenţa statistică urmează, în general, următorul algoritm:

- se obţine un eşantion printr-un procedeu valid;
- se calculează o valoare tipică a eşantionului (o statistică de sondaj);
- din considerente teoretice, sunt cunoscute repartiţia din care provine această valoare tipică şi relaţia repartiţiei de sondaj a statisticii cu valoarea tipică din populaţie; primii doi paşi pot fi interpretaţi ca şi cum s-a extras la întâmplare o valoare (valoarea calculată) dintr-o repartiţie cunoscută, cel puţin parţial (repartiţia de sondaj a statisticii);
- utilizând repartiţia de sondaj a statisticii, se pot face evaluări ale erorilor de estimaţie. În acest moment se încheie procesul prin care, plecând de la un eşantion, ajungem să facem afirmaţii în privinţa parametrilor populaţiei şi să avem o evaluare a probabilităţilor de apariţie a erorilor. *Este de remarcat faptul că nu se ştie dacă există sau nu erori, ci care sunt probabilităţile de apariţie a diverselor erori.*

Datorită importanţei deosebite care se acordă valorii medii, în continuare se particularizează acest proces pentru cazul mediei aritmetice.

Se consideră că se studiază o populaţie statistică din punctul de vedere al unei caracteristici continue X (de exemplu, înălţimea elevilor de 14 ani dintr-o zonă geografică). Să considerăm, pentru fixarea ideilor, că, în populaţia de interes, $M(X) = \mu$ şi $D^2(X) = \sigma^2$. Repartiţia variabilei X este necunoscută (deci nu sunt cunoscute μ, σ, forma distribuţiei sau numai unele dintre aceste atribute).

Din populaţia respectivă se extrage un eşantion de volum n, fie acesta $x_1, x_2, \ldots, x_n$, şi se calculează media aritmetică $\bar{x}$. Este evident că se doreşte ca această valoare să fie utilizată drept estimaţie a mediei populaţiei. Pentru a aprecia calitatea estimării, este necesar să aplicăm algoritmul anterior.

Deşi demonstraţiile nu sunt, toate, foarte complicate, dăm doar rezultatele finale asupra repartiţiei de sondaj a mediei:

$$M(\bar{x}) = \mu, \qquad D^2(\bar{x}) = \frac{\sigma^2}{n}, \qquad D(\bar{x}) = \frac{\sigma}{\sqrt{n}}.$$

Se observă astfel că $\bar{x}$ provine dintr-o repartiţie care:

- are ca valoare medie chiar media populaţiei;
- are dispersia direct proporţională cu dispersia populaţiei şi invers proporţională cu volumul eşantionului.

Forma repartiţiei este determinată de rezultatele următoare:

- dacă X este repartizată normal, atunci $\bar{x}$ este repartizată şi ea normal cu parametrii precizaţi mai sus, indiferent de volumul eşantionului;
- dacă nu se cunoaşte repartiţia lui X, atunci teorema de limită centrală afirmă că repartiţia de sondaj a lui $\bar{x}$ devine normală la limită, pentru un volum foarte mare al eşantionului. Practic, se poate accepta o repartiţie $N(\mu;\sigma^2/n)$ pentru $n > 10$, dacă repartiţia lui X este aproape simetrică, sau pentru $n > 30$ pentru repartiţii cu asimetrie pronunţată sau necunoscută.

Aceste rezultate permit acum afirmaţii probabiliste asupra mediei populaţiei. Obţinerea valorii medii $\bar{x}$ este o operaţie echivalentă cu extragerea la întâmplare a unei valori dintr-o repartiţie normală, deci probabilităţile pot fi calculate utilizând repartiţia normală. Din prezentarea repartiţiei normale, se ştie că o valoare, extrasă la întâmplare, are şansă cu atât mai mare să fie apropiată de medie, cu cât dispersia este mai mică. Prin urmare, media calculată $\bar{x}$ este cu atât mai aproape de μ, cu cât dispersia ei, σ^2/n, este mai mică. Se observă astfel că, pentru o populaţie cu o variabilitate mare, este necesar să efectuăm studii pe eşantioane mai mari (n mai mare produce micşorarea dispersiei de sondaj a mediei) decât în cazul populaţiilor cu variabilitate mai mică.

Repartiţiile de sondaj pentru alte statistici (dispersia, abaterea standard, mediana etc.) vor fi amintite doar atunci când va fi necesar. O dezbatere completă pe acest subiect poate fi urmărită în tratatele de statistică matematică.

6.7. Estimaţii

Se numeşte **estimator** orice entitate a cărei valoare poate fi utilizată drept valoare (de regulă aproximativă) pentru o altă entitate. Valoarea estimatorului poartă numele de **estimaţie**. Având ca suport situaţia în care informaţia dintr-un eşantion este utilizată pentru a obţine informaţii asupra unei populaţii, este evident că estimatorii utilizaţi în domeniul statisticii constituie valorile tipice de sondaj (cele calculate în eşantion). Valoarea care aproximează, pe baza datelor de sondaj, valoarea necunoscută a unui parametru al populaţiei poartă denumirea de **estimaţie statistică**. Astfel, media aritmetică $\bar{x}$ este estimator pentru media populaţiei μ, abaterea standard s este estimator pentru abaterea standard a populaţiei σ etc.

Din complexa teorie a estimării, prezentată pe larg în tratatele de statistică matematică, amintim aici doar unele noţiuni de bază şi câteva rezultate remarcabile. Prezentarea are şi scopul de a închega oarecum tematica acestei secţiuni. O atenţie specială este acordată totuşi intervalelor de încredere.

Dacă se notează cu δ_n^* estimatorul pentru parametrul δ, unde n semnifică volumul eşantionului, se spune că δ_n^* este **nedeplasat** dacă

$$M(\delta_n^*) = \delta,$$

adică media repartiţiei de sondaj a estimatorului este egală cu parametrul estimat.

Estimaţiile pentru care

$$M(\delta_n^*) = \delta + a_n \text{ sau } M(\delta_n^*) = \delta - a_n,$$

unde $a_n \to 0$ pentru $n \to \infty$, se numesc estimaţii **deplasate pozitiv**, respectiv **deplasate negativ**. Cu alte cuvinte, are loc, în special pentru eşantioane mici, o supraevaluare, respectiv o subevaluare, a valorii parametrului.

Un estimator este **absolut corect** dacă

$$M(\delta_n^*) = \delta \text{ şi } \lim_{n\to\infty} D^2(\delta_n^*) = 0,$$

altfel spus dacă, o dată cu creşterea volumului eşantionului, are loc şi o micşorare a dispersiei estimatorului.

Estimatorul nedeplasat care are dispersia minimă pentru un volum dat al eşantionului poartă denumirea de **estimaţie eficientă**. Minimumul este selectat dintre toţi estimatorii posibili pentru parametrul implicat.

În ceea ce priveşte statisticile uzuale, se demonstrează că:

- $\bar{x}$ este o estimaţie absolut corectă şi eficientă a mediei μ a populaţiei;
- s^2 (calculat cu numitorul n – 1) este o estimaţie absolut corectă a dispersiei σ^2 a populaţiei; s^2 (calculată cu numitorul n) nu este nedeplasată;
- abaterea standard nu este nedeplasată;
- pentru repartiţia binomială Bi(n;p), o estimaţie eficientă a parametrului p este frecvenţa relativă a cazurilor „succes".

După natura lor, în statistică se utilizează două tipuri de estimaţii:

- punctuale;
- sub formă de interval.

Ambele categorii de estimaţii se înscriu pe linia extinderii simţului comun prin raţionamentul statistic. Astfel, în mod uzual, se întâlnesc în viaţa cotidiană două moduri de precizare numerică a unei situaţii: sau indicarea unei valori (*afară sunt 5 grade*), sau indicarea unui interval (*afară sunt între 3 şi 6 grade*).

6.7.1. Estimaţii punctuale

Prin **estimaţie punctuală** se înţelege valoarea unui estimator calculată într-un eşantion. Astfel, media aritmetică obţinută într-un eşantion este o estimaţie punctuală a mediei populaţiei.

Numim **eroare de estimare** valoarea absolută a diferenţei dintre estimaţia punctuală şi valoarea parametrului, $|\delta_n^* - \delta|$, unde δ_n^* este estimaţia punctuală, adică valoarea estimatorului obţinută într-un eşantion de volum n, iar δ este valoarea parametrului corespunzător din populaţie.

Deoarece nu se cunoaşte valoarea parametrului, este practic imposibil să se calculeze eroarea de estimare. Pentru estimatori particulari se pot obţine judecăţi probabiliste pe baza repartiţiei de sondaj. Astfel, pentru media aritmetică, s-a precizat că repartiţia de sondaj poate fi considerată normală pentru $n > 30$, $N(\mu;\sigma^2/n)$. Prin urmare, din proprietăţile repartiţiei normale rezultă că media este la distanţa de o abatere standard în 68% dintre cazuri, la două abateri standard în aproximativ 95% dintre cazuri etc. Deci eroarea de estimare este de cel mult $\sigma/\sqrt{n}$ cu probabilitatea de 68%, este de cel mult $2\sigma/\sqrt{n}$ cu probabilitatea de 95% etc.

De exemplu, să presupunem că pentru o populaţie statistică ştim $\sigma = 12$ şi dintr-un eşantion de volum n = 36 am calculat $\bar{x} = 42$. Obţinem

$$D(\bar{x}) = \frac{\sigma}{\sqrt{n}} = \frac{12}{\sqrt{36}} = \frac{12}{6} = 2.$$

Putem interpreta acum: există 68% şanse ca $\bar{x} = 42$ să difere de μ cu cel mult 2, există 95% şanse ca $\bar{x} = 42$ să difere de μ cu cel mult 4 etc.

Să presupunem că $\sigma = 24$, restul datelor fiind neschimbate. Se obţine $D(\bar{x}) = 4$. Putem spune că există 95% şanse ca media calculată, 42, să difere de μ cu cel mult 8. Se consideră că, în acest caz, estimarea este mai puţin precisă.

Se susţine, încă o dată, din acest exemplu că, pentru o abatere standard mică $D(\bar{x})$, *estimarea va fi mai precisă, în sensul că erorile mici vor fi obţinute cu probabilităţi mari* (prin compensaţie, erorile mari se vor obţine cu probabilităţi mici).

Datorită rolului evidenţiat în paragraful anterior, cantitatea $D(\bar{x})$ mai poartă denumirea de **eroare standard a mediei.**

6.7.2. Intervale de încredere

Deoarece aprecierea erorilor de estimare posibile în cazul estimaţiilor punctuale este suficient de laborioasă, s-a încercat dezvoltarea unei metode care să pună accentul pe probabilităţile de apariţie a erorilor. Această metodă generică este cea a determinării intervalelor de încredere sau metoda de estimare sub formă de interval.

Fie o populaţie statistică, ce este caracterizată de o v.a. continuă X a cărei repartiţie depinde de un parametru δ, necunoscut. Problema propusă este ca, având la dispoziţie datele de sondaj ale unui eşantion de volum n, să se determine două limite între care valoarea parametrului δ să se afle cu o probabilitate fixată. Cu alte cuvinte, să se determine două statistici, δ_1 şi δ_2, astfel încât

$$P(\delta_1 < \delta < \delta_2) = p,$$

unde p este o probabilitate, independentă de parametrul δ. Explicit, relaţia exprimă faptul că valoarea parametrului este situată între cele două valori, δ_1 şi δ_2, cu probabilitatea p.

Prin definiţie, dacă se pot determina δ_1 şi δ_2 astfel încât pentru o valoare α prestabilită ($0 < \alpha < 1$) să aibă loc

$$P(\delta_1 < \delta < \delta_2) = 1 - \alpha,$$

atunci intervalul (δ_1, δ_2) se numeşte **interval de încredere** pentru parametrul necunoscut δ cu un **coeficient** (sau **nivel**) **de încredere** egal cu $1 - \alpha$ sau cu o **siguranţă statistică** $S_\alpha = 1 - \alpha$.

Valoarea α se numeşte **nivel** (sau **prag**) **de semnificaţie.** Utilizarea termenului *semnificaţie* este explicată în secţiunea dedicată testelor statistice.

Din punctul de vedere al estimaţiei, eroarea de estimare este egală, cel mult, cu jumătatea intervalului de încredere. Este de dorit astfel ca intervalul de încredere să aibă o lungime cât mai mică pentru ca estimarea să fie mai precisă.

Dacă valoarea parametrului δ este fixată o dată cu populaţia, limitele intervalului de încredere depind de eşantionul ales, deci sunt limite aleatorii. Se va subînţelege atunci că *nu putem avea certitudinea* că valoarea parametrului δ se situează în intervalul determinat. Raţionamentul corect este acela că, dacă se repetă procedeul pentru eşantioane diferite, atunci $100(1 - \alpha)\%$ dintre intervalele determinate vor conţine valoarea necunoscută a parametrului δ (evident că nu se va şti care sunt acele intervale). Apare atunci necesitatea ca α să fie cât mai mic, dar, o dată cu micşorarea lui α, se măreşte lungimea intervalului, deci erorile de estimare sunt mai mari.

În mod uzual se alege $\alpha = 0{,}05$ sau $\alpha = 0{,}01$.

Dacă atât δ_1, cât şi δ_2 sunt finite, atunci intervalul de încredere se numeşte **bilateral.** În cazul în care δ_1 este $-\infty$ sau δ_2 este $+\infty$, ceea ce duce de fapt la determinarea unei singure limite, intervalul se numeşte **unilateral.**

Relaţia dintre nivelul de încredere şi lungimea intervalului este evidentă chiar la nivelul simţului comun. O afirmaţie de genul „afară sunt între 10 şi 20 de grade" are şanse mai mari să fie mai apropiată de adevăr decât o afirmaţie de genul „afară sunt între 14 şi 16 grade". Este însă evident că a doua afirmaţie, dacă este adevărată, este mai precisă. Deci precizia şi încrederea sunt două noţiuni oarecum contrare şi compromisul între ele conduce la o stare acceptabilă statistic.

În continuare sunt prezentate intervalele de încredere pentru statisticile uzuale: parametrii repartiţiei normale (media, dispersia) şi parametrul repartiţiei binomiale.

6.7.2.1. Intervale de încredere pentru valoarea medie

Fie o populaţie statistică ce este caracterizată de o v.a. X repartizată normal, cu parametrii μ şi σ^2, $X \sim N(\mu, \sigma^2)$. Presupunem că în urma unui sondaj s-au obţinut, dintr-un eşantion de volum n, media de sondaj $\bar{x}$ şi dispersia de sondaj s^2 (calculată cu numitorul $n - 1$). Fixăm un prag de semnificaţie α.

Determinarea unui interval de încredere pentru media populaţiei presupune determinarea a două limite m_1 şi m_2, astfel încât:

$$P(m_1 < \mu < m_2) = 1 - \alpha.$$

Limitele m_1 şi m_2 depind atât de eşantion, cât şi de cunoaşterea sau nu a dispersiei populaţiei.

Dispersia, σ^2, este cunoscută.

Din repartiţia de sondaj a mediei rezultă atunci că intervalul de încredere pentru medie, cu siguranţa statistică $S_\alpha = 1 - \alpha$, este

$$\bar{x} - \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2} < \mu < \bar{x} + \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2},$$

unde $z_{1-\alpha/2}$ este quantila de ordinul $1 - \alpha/2$ a variabilei normale reduse.

Intervalul definit,

$$\left(\bar{x} - \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2}, \bar{x} + \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2} \right),$$

are înţelesul că, dacă se repetă pentru o sută de eşantioane de acelaşi volum n, atunci în $100(1 - \alpha/2)$ dintre cazuri ne putem aştepta ca intervalul să conţină valoarea exactă a mediei populaţiei.

Diferenţa maximă dintre media populaţiei şi media de sondaj este

$$\varepsilon = \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2},$$

care reprezintă eroarea de estimare.

Dacă se rezolvă ecuaţia anterioară pentru n, se obţine

$$n = \left(\frac{\sigma \cdot z_{1-\alpha/2}}{\varepsilon} \right)^2,$$

care se poate utiliza pentru determinarea volumului eşantionului necesar pentru a obţine o mărime fixată a erorii. Se observă că n este invers proporţional cu pătratul erorii ε; aşadar, pentru a înjumătăţi eroarea ar trebui să mărim de patru ori volumul eşantionului.

În afara intervalului bilateral, sunt uneori utile intervale unilaterale definite prin:

$$P(\mu < \bar{x} + \frac{\sigma}{\sqrt{n}} z_{1-\alpha}) = 1 - \alpha,$$

$$P(\mu > \bar{x} - \frac{\sigma}{\sqrt{n}} z_{1-\alpha}) = 1 - \alpha,$$

care arată, cu aceeaşi siguranţă statistică, unde se situează media populaţiei în raport cu media de sondaj. Intervalele sunt aici nelimitate într-un sens.

Dispersia, σ^2, este necunoscută.

Dacă nu se cunoaşte valoarea dispersiei populaţiei, σ^2, atunci aceasta este estimată prin dispersia de sondaj, s^2.

Se demonstrează că intervalul bilateral de încredere este dat de

$$\bar{x} - \frac{s}{\sqrt{n}} t_{1-\alpha/2;\nu} < \mu < \bar{x} + \frac{s}{\sqrt{n}} t_{1-\alpha/2;\nu},$$

unde prin $t_{1-\alpha/2;\nu}$ se notează quantila de ordin 1-α/2 a distribuţiei t cu $\nu = n - 1$ grade de libertate.

Intervalele unilaterale, la acelaşi prag de semnificaţie, sunt:

$$\mu < \bar{x} + \frac{s}{\sqrt{n}} t_{1-\alpha;\nu},$$

$$\mu > \bar{x} - \frac{s}{\sqrt{n}} t_{1-\alpha;\nu}.$$

Pentru $n > 30$, se pot utiliza aproximările $t_{1-\alpha/2;\nu} \cong z_{1-\alpha/2}$ şi $t_{1-\alpha;\nu} \cong z_{1-\alpha}$, formulele obţinute fiind aplicabile în cazul eşantioanelor mari.

Deoarece quantilele distribuţiei t depind şi ele de n, nu se poate determina volumul eşantionului, pentru a obţine o eroare fixată, decât prin încercări.

6.7.2.2. Intervale de încredere pentru dispersie

Fie o populaţie normală sau aproximativ normală, cu parametrii μ şi σ^2 necunoscuţi. Se demonstrează că intervalul de încredere bilateral pentru dispersia populaţiei, cu încrederea statistică de $1 - \alpha$, este dat de

$$\frac{(n-1)s^2}{\chi^2_{1-\alpha/2;\nu}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{\alpha/2;\nu}},$$

unde n este volumul eşantionului, s^2 este dispersia de sondaj, iar $\chi_{\alpha/2;\nu}$ şi $\chi_{1-\alpha/2;\nu}$ sunt quantilele de ordin α/2, respectiv $1 - \alpha/2$, ale repartiţiei χ^2 cu $\nu = n - 1$ grade de libertate.

Intervalele de încredere unilaterale sunt date de

$$\sigma^2 < \frac{(n-1)s^2}{\chi^2_{\alpha;\nu}},$$

$$\sigma^2 > \frac{(n-1)s^2}{\chi^2_{1-\alpha;\nu}}.$$

În calculele efective, pentru $n > 30$, se poate recurge la aproximarea quantilelor repartiţiei χ^2 cu quantilele repartiţiei normale standard, N(0;1), dată de relaţia:

$$\chi^2_{p;\nu} \cong \frac{1}{2}\left[z_p + \sqrt{2\nu - 1}\right]^2,$$

unde p este ordinul necesar al quantilei (α, $1 - \alpha$ etc.).

6.7.2.3. Intervale de încredere pentru abaterea standard

Intervalele de încredere pentru abaterea standard, σ, se pot calcula după formulele:

$$\frac{s}{1 + \frac{z_{1-\alpha/2}}{\sqrt{2n}}} < \sigma < \frac{s}{1 - \frac{z_{1-\alpha/2}}{\sqrt{2n}}}$$

pentru cazul bilateral şi

$$\sigma < \frac{s}{1 - \frac{z_{1-\alpha}}{\sqrt{2n}}},$$

$$\sigma > \frac{s}{1 + \frac{z_{1-\alpha}}{\sqrt{2n}}},$$

pentru intervalele unilaterale.

Siguranţa statistică este $S_\alpha = 1-\alpha$.

În aplicaţiile în care nu se urmăreşte un mare grad de precizie, se pot utiliza, pentru limitele de încredere ale abaterii standard, şi valorile obţinute prin extragerea rădăcinii pătrate din limitele de încredere pentru dispersie (mai ales dacă acestea au fost calculate anterior).

Determinarea unui volum n al eşantionului care să permită o eroare de estimare impusă se poate face doar prin încercări, atât la dispersie, cât şi la abaterea standard.

6.7.2.4. Intervale de încredere pentru diferenţa mediilor a două populaţii normale

Deoarece multe cercetări practice presupun compararea a două metode, două grupuri etc., este de un deosebit interes să se determine un interval de încredere pentru diferenţa observată.

Pentru fixarea ideilor, fie două populaţii caracterizate de variabilele aleatorii X_1 şi, respectiv, X_2, independente, ambele repartizate normal

$$X_1 \sim N(\mu_1;\sigma_1^2),\ X_2 \sim N(\mu_2;\sigma_2^2).$$

Din prima populaţie dispunem de un eşantion de volum n_1, iar din a doua populaţie dispunem de un eşantion de volum n_2.

Valorile tipice calculate în cele două eşantioane sunt $\bar{x}_1, s_1^2$ şi, respectiv, $\bar{x}_2, s_2^2$.

Deoarece este evident că mediile de sondaj pot să difere între ele atât din cauza sondajului, cât şi datorită diferenţei dintre mediile populaţiilor, este important ca, pornind de la diferenţa observată, să se poată stabili intervale de încredere pentru diferenţa mediilor populaţiilor, $\mu_1 - \mu_2$. Pentru aplicaţii, este de la sine înţeles că variabilele X_1 şi X_2 sunt comparabile (mai precis, aceeaşi caracteristică urmărită în două grupuri etc.).

Intervalele de încredere au expresii diferite, după cum dispersiile sunt cunoscute sau necunoscute, sunt egale sau inegale. Pentru a stabili egalitatea dispersiilor celor două populaţii, se va aplica un test statistic de comparare (vezi capitolul dedicat testării ipotezelor statistice).

Dispersiile σ_1^2, σ_2^2 sunt cunoscute.

În acest caz, intervalul de încredere bilateral, cu siguranţa statistică $1 - \alpha$, este dat de

$$(\bar{x}_1 - \bar{x}_2) - z_{1-\alpha/2} \cdot \sigma_d < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + z_{1-\alpha/2} \cdot \sigma_d,$$

unde σ_d este abaterea standard a diferenţei mediilor de sondaj, dată de

$$\sigma_d = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}.$$

Intervalele de încredere unilaterale sunt, cu aceleaşi notaţii,

$$\mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + z_{1-\alpha} \cdot \sigma_d,$$
$$\mu_1 - \mu_2 > (\bar{x}_1 - \bar{x}_2) - z_{1-\alpha} \cdot \sigma_d.$$

Dispersiile σ_1^2, σ_2^2 sunt necunoscute, dar egale.

Presupunem că σ_1^2, σ_2^2 sunt egale (testarea ipotezei impune aplicarea unui test F), adică $\sigma_1^2 = \sigma_2^2 = \sigma^2$, unde σ^2 este notaţia pentru valoarea comună a dispersiilor. În acest caz, se estimează pentru început dispersia comună prin

$$s^2 = \frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 + n_2 - 2}.$$

Se demonstrează că intervalul de încredere bilateral, la nivelul de semnificaţie α, este

$$(\bar{x}_1 - \bar{x}_2) - t_{1-\alpha/2;\nu} \cdot s_d < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{1-\alpha/2;\nu} \cdot s_d,$$

unde $\nu = n_1 + n_2 - 2$ este numărul gradelor de libertate a diferenţei de sondaj a mediei, $t_{1-\alpha/2;\nu}$ este quantila repartiţiei t cu ν grade de libertate, iar s_d este abaterea standard a diferenţei:

$$s_d = s \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}.$$

Intervalele unilaterale, utilizând aceleaşi notaţii, sunt date de:

$$\mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{1-\alpha;\nu} \cdot s_d,$$
$$\mu_1 - \mu_2 > (\bar{x}_1 - \bar{x}_2) - t_{1-\alpha;\nu} \cdot s_d.$$

Intervalele de încredere specificate rămân acceptabile şi pentru repartiţii care nu se abat prea mult de la cea normală, iar n_1 şi n_2 sunt cel puţin egale cu 20.

Dispersiile σ_1^2, σ_2^2 sunt necunoscute şi inegale.

Atunci când ipoteza egalităţii dispersiilor nu poate fi acceptată, se demonstrează că intervalele de încredere a diferenţei mediilor populaţiilor sunt date de formulele:

$$(\bar{x}_1 - \bar{x}_2) - t_{1-\alpha/2;\nu} \cdot s_a < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{1-\alpha/2;\nu} \cdot s_a,$$

$$\mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{1-\alpha;\nu} \cdot s_a,$$
$$\mu_1 - \mu_2 > (\bar{x}_1 - \bar{x}_2) - t_{1-\alpha;\nu} \cdot s_a,$$

unde ν, numărul gradelor de libertate pentru repartiţia Student asociată, este dat de relaţia

$$\nu = \frac{(n_1 - 1)(n_2 - 1)}{(n_2 - 1)c^2 + (n_1 - 1)(1 - c)^2} \text{ în care } c = \frac{\frac{s_1^2}{n_1}}{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}.$$

Abaterea standard s_a a diferenței mediilor de sondaj este dată de

$$s_a = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}},$$

cu aceasta explicându-se toate entitățile din formulele intervalelor de încredere.

6.7.2.5. *Intervale de încredere pentru un raport*

Este frecventă situația în care, pornind de la un eșantion de volum n, cercetătorul este interesat de frecvența relativă (exprimată, eventual, procentual) a cazurilor care prezintă o anumită caracteristică. De exemplu, într-un eșantion aleatoriu de 50 de elevi sunt 20 (deci 40%) care doresc să-și continue studiile. Apare întrebarea care este procentajul celor care doresc același lucru în întreaga populație școlară, deci o problemă care se poate rezolva prin determinarea unui interval de încredere pentru raportul (procentajul) celor din populația școlară care doresc să-și continue studiile.

Situația poate fi formalizată prin luarea în considerare a unei v.a. X, repartizată binomial, Bi(n;p), pentru care se consideră valoare „succes" prezența caracteristicii dorite. Se știe că frecvența relativă, notată aici f_n, unde indicele n este atașat pentru a evidenția volumul eșantionului, este o estimație eficientă a parametrului p.

Dintre situațiile de definire a intervalelor de încredere pentru parametrul p, luăm în considerare doar cazul eșantioanelor mari sau, mai precis, cazul în care se poate utiliza aproximarea repartiției binomiale prin repartiția normală (a se vedea discuția de la prezentarea repartiției binomiale). În acest caz, se demonstrează că intervalul de încredere cu siguranța statistică $S_\alpha = 1 - \alpha$ este dat de

$$f_n - z_{1-\alpha/2} \cdot \sqrt{\frac{f_n(1-f_n)}{n}} < p < f_n + z_{1-\alpha/2} \cdot \sqrt{\frac{f_n(1-f_n)}{n}},$$

unde f_n este frecvența relativă a cazurilor „succes" din eșantionul de volum n, iar $z_{1-\alpha/2}$ este quantila de ordin $(1-\alpha/2)$ a repartiției normale standard.

Pentru exemplul considerat, fixând $\alpha = 0{,}05$, au loc relațiile:

$$f_n = 20/50 = 0{,}40\,;\ 1-f_n = 30/50 = 0{,}60\,;\ z_{1-\alpha/2} = 1{,}96$$

$$0{,}40 - 1{,}96 \times \sqrt{\frac{0{,}40 \times 0{,}60}{50}} < p < 0{,}40 + 1{,}96 \times \sqrt{\frac{0{,}40 \times 0{,}60}{50}},$$

adică

$$0{,}264 < p < 0{,}536.$$

Concluzia este că ne putem aștepta ca procentajul elevilor care doresc să-și continue studiile să fie între 26,4% și 53,6%, cu o încredere statistică de 95%.

Capitolul 7

TESTAREA IPOTEZELOR STATISTICE

Reamintim că am definit statistica drept un corp de metode utilizate pentru a aduna, a descrie și a analiza date numerice. Se poate crea astfel convingerea că atitudinea favorită a statisticianului este aceea de „detectiv" (Tukey, 1977), deci atitudinea prin care cercetătorul, cufundat într-un ocean mai mare sau mai mic de date numerice, încearcă să găsească, să descopere, prin analiza acestora, legități, relații, structurări ale realității studiate.

O asemenea atitudine este greșită, chiar dacă este favorizată în ultimul timp, în special datorită mijloacelor electronice de stocare și prelucrare a datelor.

De fapt, datele nu sunt culese la întâmplare măcar pentru motivul că obținerea și prelucrarea lor este uneori foarte costisitoare. Fără a încerca o generalizare de dragul generalizării, se poate accepta ideea că, în cele mai multe prelucrări statistice, datele sunt obținute și prelucrate pentru a verifica ipoteze ale cercetătorilor (potrivit demersului științific general). Deci, ca o primă imagine a capitolului de față, trebuie reținută secvența:

- formularea unei ipoteze;
- obținerea de date experimentale;
- verificarea ipotezei pe baza acestor date.

Deși, aparent, obiectul verificării ipotezelor se regăsește doar în a treia afirmație, se va vedea că nu se pot separa cele trei faze enunțate.

Analiza secvenței anterioare arată, pentru început, că problematica se înscrie tot în inferența statistică, deoarece ipotezele importante trebuie să privească populațiile statistice, datele experimentale presupun eșantioane reprezentative, iar verificarea trebuie să conducă la afirmații privitoare, din nou, la populații.

7.1. *Ipoteze statistice*

Prin **ipoteză**, dicționarele definesc în mod uzual o presupunere, o explicație provizorie, enunțată pe baza unor fapte cunoscute, cu privire la esența, cauza, legea, mecanismul intern al unui fenomen, și care este supusă verificării sau demonstrării. Se va utiliza noțiunea de **ipoteză statistică**, deoarece verificarea se face pe baza evidenței statistice, deci pe baza informațiilor obținute, în general, din eșantioane.

Datorită educației deterministe, ne așteptăm ca testarea, verificarea unei ipoteze să fie o procedură riguroasă în sensul că, în final, putem spune că ipoteza este falsă sau că este adevărată. Dar, din motive metodologice valide, procesul verificării este mai complicat. El se bazează pe *criteriul falsificabilității* (concept lansat de Karl Popper în *Logica cercetării*) care afirmă că, în timp ce este posibil să determinăm când o ipoteză este falsă, este mult mai dificil, dacă nu imposibil, să demonstrăm că o ipoteză este adevărată.

Dacă evidența empirică, a datelor, este contrară ipotezei, ipoteza este falsă. Dar, dacă evidența coincide cu ipoteza, nu rezultă în mod necesar că ipoteza este adevărată. În acest caz, singurul lucru rezonabil care se poate afirma este acela că evidența disponibilă nu a arătat încă falsitatea ipotezei.

Următorul exemplu (K. Popper) este sugestiv. Ipoteză: *toate lebedele sunt albe*. Pentru a verifica ipoteza, este natural să se caute lebede. Găsirea oricâtor lebede albe nu arată decât că afirmația nu poate fi respinsă (dar nu că este adevărată). Urmând demersul inductiv, considerăm că lebedele sunt albe, deoarece nu am găsit nici una neagră. Găsirea unei singure lebede negre demonstrează însă falsitatea ipotezei.

Să reținem din acest exemplu că, în ipoteza formulată, doar găsirea unei lebede negre este un eveniment semnificativ, fiind singurul care conduce la o concluzie precisă. Această noțiune, foarte des utilizată în cercetările statistice, este păstrată și în continuare, adică vom considera **semnificativ** un eveniment care contrazice ipoteza de plecare. Sensul se apropie de cel curent prin aceea că, falsificând ipoteza de pornire, se permite acceptarea unei alte ipoteze, în acest fel dezvoltându-se procesul de cunoaștere.

7.2. Teste statistice

Pentru a evidenția mecanismul unui test statistic, se prezintă, în continuare, două exemple tipice: un exemplu dezvoltat de R.A. Fisher (1935), care oferă întregul suport intuitiv necesar înțelegerii unor alte proceduri, și un test legat de media aritmetică.

Exemplul 7.1. (Fisher). O persoană pretinde că poate recunoaște dacă într-un pahar de ceai cu lapte s-a turnat mai întâi laptele (ne aflăm în Anglia, unde problema este importantă atât pentru subiect – ceai –, cât și pentru pariurile care se pot încheia). Persoanei respective i se prezintă, într-o ordine întâmplătoare, 8 pahare, cerându-i-se să desemneze cele 4 pahare în care laptele a fost turnat mai întâi. Este de la sine înțeles că paharele nu se deosebeau prin nimic unele de altele. În urma alegerii, persoana indică în mod corect 3 pahare, al patrulea fiind indicat eronat. Trebuie să decidem dacă are sau nu capacitatea asumată.

Este evident că un rezultat izolat nu poate demonstra acest lucru cu certitudine, dar se poate dezvolta o procedură de decizie astfel încât să conducă la un răspuns fals cu o probabilitate mică. În procesul de decizie trebuie să se discearnă, pe cât

posibil, între un rezultat bun care este obținut întâmplător (în definitiv, sunt câștigători la loterii care joacă o singură dată, din întâmplare) și un rezultat bun datorat capacității asumate.

Se începe prin enunțarea a două ipoteze:

H_0: *subiectul nu poate distinge dacă laptele a fost turnat mai întâi;*

H_1: *subiectul poate distinge dacă laptele a fost turnat mai întâi.*

Ipoteza H_0, numită în mod obișnuit **ipoteza nulă**, este cea care stă la baza întregului raționament ulterior. Ipoteza H_1, numită **ipoteza alternativă**, este contrara ipotezei nule. Fixarea acestor ipoteze nu se face la întâmplare; explicații ulterioare vor completa dezbaterea.

Situându-ne în ipoteza H_0, rezultă că subiectul alege la întâmplare 4 pahare dintre cele 8 prezentate.

Cum există 70 de modalități de alegere a patru pahare din opt ($= C_8^4$), rezultă că subiectul a avut de ales între 70 de variante. Șansa de a alege o variantă este de 1/70 (variante echiprobabile) datorită ipotezei H_0 pe care fundamentăm raționamentul. Notând cu „c" un pahar indicat corect și cu „f" un pahar indicat eronat, rezultatele posibile ale experimentului sunt:

Rezultate posibile	Frecvență	Probabilitate (sub H_0)
4 c + 0 f	1	1/70 = 0,014
3 c + 1 f	16	16/70 = 0,229
2 c + 2 f	36	36/70 = 0,514
1 c + 3 f	16	16/70 = 0,229
0 c + 4 f	1	1/70 = 0,014

În calculul frecvențelor se observă simetria rezultatelor, valorile fiind obținute după procedeele uzuale din analiza combinatorie. De exemplu, numărul de combinații de tip 3c + 1f este determinat prin produsul dintre alegerile a 3 pahare corect indicate (din 4 existente) și alegerile unui pahar greșit indicat (din 4 existente):

$$C_4^3 \times C_4^1 = 4 \times 4 = 16$$

În mod analog, numărul de combinații 2c + 2f este $C_4^2 \times C_4^2 = 6 \times 6 = 36$

În ipoteza H_0 toate combinațiile sunt egal probabile, deci pentru calculul probabilităților se poate efectua raportul dintre numărul cazurilor favorabile și numărul total de cazuri.

Fiecare linie din tabel reprezintă un eveniment elementar al probei. S-a evidențiat linia reprezentând evenimentul elementar care a avut loc (subiectul a ghicit 3c + 1f).

Cu aceste evenimente elementare se formează două evenimente compuse în modul următor:

– C, cuprinzând evenimentul realizat și toate evenimentele extreme în raport cu acesta și conducând la respingerea ipotezei H_0. În cazul concret, în C, pe

lângă evenimentul realizat 3c + 1f, mai intră evenimentul 4c + 0f, adică evenimentul ce reprezintă ghicirea completă a celor patru pahare în care laptele a fost turnat mai întâi.

- $\overline{C}$, evenimentul contrar lui C. În exemplu, $\overline{C}$ conține evenimentele 2c + 2f, 1c + 3f şi 0c + 4f.

Din regulile de calcul cu probabilități rezultă că P(C) = 0,243 şi, prin urmare, $P(\overline{C}) = 1 - P(C) = 0,757$.

Cele două principii esențiale care au condus procesul până aici sunt următoarele:

- se poate calcula probabilitatea lui C sub H_0; evenimentele elementare sunt egal probabile doar în condițiile ipotezei nule. Posibilitatea respectivă este, de altfel, motivul pentru care s-a ales această ipoteză nulă, şi nu alta.
- probabilitatea lui C este mai mare sub H_1 decât sub H_0; faptul este evident pentru că, dacă subiectul are capacitatea de a discerne ordinea turnării lichidelor în pahar, va efectua de regulă alegeri corecte. Trebuie însă menționat faptul că nu putem şti în acest caz probabilitatea lui C, ea fiind funcție de gradul de posesie a capacității asumate.

Numim **probabilitate critică** a testului, notată p_c, probabilitatea evenimentului C sub ipoteza H_0:

$$p_c = P(C) = P_{H_0}(C).$$

Regula de decizie în test este că, dacă această probabilitate este considerată „mică", atunci se respinge ipoteza nulă. Într-adevăr, evenimentul C, care s-a realizat (el conține rezultatul observat), este considerat „miraculos", total neașteptat în contextul ipotezei H_0, în comparație cu contextul ipotezei H_1, care-i asigură o probabilitate mult mai mare de apariție.

Dacă însă probabilitatea critică nu este considerată prea „mică", nu putem respinge ipoteza H_0. Este ca şi cum a avut loc un eveniment cu privire la care ne aşteptam să se întâmple; deci nu s-a petrecut nimic semnificativ, care să ne schimbe baza raționamentului.

În exemplul dat, probabilitatea critică de 0,243 nu poate fi considerată mică (e adevărat că o asemenea şansă la loteria națională ar asigura, în medie, un câştig la fiecare patru jocuri). Prin urmare, nu se poate respinge ipoteza H_0, că persoana implicată nu are capacitatea de a distinge ordinea turnării lichidelor în pahare.

Rămâne o singură întrebare: *începând de unde o probabilitate este considerată „mică"*? Pentru a nu introduce subiectivismul în această decizie, foarte importantă pentru desfăşurarea unui test statistic, se fixează, anterior deciziei în test, un prag sub care o probabilitate este considerată „mică". Această valoare se numeşte **prag de semnificație** şi se notează în mod obişnuit cu α. Denumirea este justificată pentru că desparte evenimentele semnificative (care conduc la respingerea ipotezei nule) de cele nesemnificative (care nu permit respingerea ipotezei nule).

Valorile uzuale pentru α sunt 0,05 sau 0,01, valori stabilite atât prin convenție, cât şi datorită unor considerente teoretice. Analizând procesul de decizie prezentat,

se poate observa că fixarea pragului de semnificație α, de exemplu $\alpha = 0,05 = 5\%$, este echivalentă cu acceptarea riscului de a respinge în mod fals ipoteza H_0 în 5% din cazuri; cu alte cuvinte, putem greşi în 5% din situațiile în care respingem ipoteza nulă. În exemplul dat, p_c este deasupra oricăreia dintre valorile α uzuale, deci este mare şi, prin urmare, nu se respinge ipoteza nulă.

Cu alte cuvinte, regula de decizie în test poate fi formulată:

- dacă $p_c \le \alpha$, atunci respingem ipoteza nulă, H_0, în favoarea ipotezei alternative, H_1;
- dacă $p_c > \alpha$, atunci nu respingem ipoteza nulă H_0.

De reținut exprimarea „nu respingem H_0" în loc de „acceptăm H_0", ultima având mai mult sensul atribuirii valorii de adevăr lui H_0, sens care am convenit că nu este corect. Se poate utiliza şi formularea „tolerăm H_0", care lasă loc unei urme de îndoială, atitudine corectă în sensul dezbaterii de până acum.

Dacă α se fixează înainte de decizia în test, se poate analiza formarea unui eveniment C_α maxim (în sensul analogiei cu teoria mulțimilor), printr-o reunire de evenimente elementare, astfel încât $P(C_\alpha) \le \alpha$. Evident, formarea lui C_α respectă principiile enunțate mai sus, adică se reunesc acele evenimente elementare ce sunt mai probabile sub ipoteza H_1 decât sub ipoteza nulă. În urma efectuării experimentului, se cercetează dacă evenimentul produs aparține sau nu lui C_α. În cazul în care evenimentul produs este în C_α, atunci ipoteza nulă se respinge, întrucât probabilitatea critică satisface, prin construcția făcută, relația

$$p_c \le P(C_\alpha) \le \alpha.$$

Imaginea următoare trebuie înțeleasă dinamic:

1) se fixează α;
2) se stabileşte regiunea C_α;
3) se obține rezultatul experimentului;
4) se decide în test după regiunea în care se încadrează rezultatul.

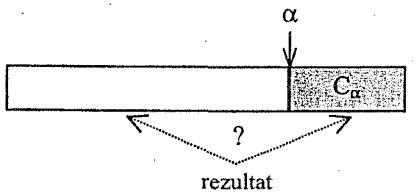


Figura 7.1. *Regiunea de respingere*

Se observă astfel că valoarea fixată pentru α împarte mulțimea rezultatelor posibile într-o mulțime de rezultate care conduc la respingerea ipotezei nule şi o mulțime de rezultate care conduc la tolerarea ipotezei nule.

Numim **regiune de respingere**, pentru un nivel de semnificaţie α fixat, mulţimea rezultatelor probei care conduc la respingerea ipotezei H_0. Dacă se pot defini limitele numerice ale regiunii de respingere, acestea vor fi numite, uneori, **valori critice** ale testului.

În exemplul dat, pentru $\alpha = 0{,}05$, se obţine ca regiune de respingere doar rezultatul „4c + 0f", adică subiectul ar fi trebuit să aleagă corect toate cele patru pahare. Rezultatul observat, „3c + 1f", nu se încadrează în regiunea de respingere; prin urmare, nu putem respinge ipoteza nulă la nivelul $\alpha = 0{,}05$.

În cazul în care se fixează $\alpha = 0{,}01$, regiunea de respingere este mulţimea vidă (evenimentul situat extrem, 4c + 0f, are singur probabilitatea mai mare). Semnificaţia practică este aceea că experimentul nu poate discerne cu încrederea statistică de 99% între cele două ipoteze. Ar trebui să se schimbe în acest scop condiţiile cercetării: de exemplu, să se propună spre alegere un număr mai mare de pahare etc.

Sintetizând ideile expuse, se poate stabili că procedura de testare a unei ipoteze statistice (pe scurt, un **test statistic**) constă în:

1. formularea unei ipoteze H_0 şi a unei ipoteze alternative H_1;
2. obţinerea de date statistice;
3. calcularea unei statistici din datele experimentale, statistică pentru care se ştie repartiţia în ipoteza H_0;
4. decizia pe baza probabilităţii critice şi a nivelului de semnificaţie α sau decizia pe baza regiunii de respingere corespunzătoare nivelului α fixat.

Particularitatea exemplului luat în considerare constă în aceea că s-a calculat repartiţia statisticii şi nu s-a făcut nici o presupunere apriorică asupra acestei distribuţii şi a probabilităţilor asociate. Acest aspect înscrie testul efectuat în categoria **testelor neparametrice**. Datorită extinderii puterii de calcul electronic, dezvoltarea unor astfel de teste este din ce în ce mai accentuată şi, practic, se poate imagina o procedură asemănătoare pentru multe ipoteze de interes.

Exemplul 7.2. Să presupunem că se studiază o colectivitate umană adultă despre care se cunoaşte dintr-un studiu anterior că are înălţimea medie de $\mu = 1{,}68$ m şi o abatere standard $\sigma = 0{,}10$ m. Este vorba de un studiu longitudinal, care determină valorile parametrilor la anumite intervale de timp. Se pune problema de a şti dacă înălţimea medie s-a modificat faţă de determinarea anterioară, cauzele acestei modificări nu sunt discutate în exemplul dat.

O ipoteză tipică este:

$$H_0: \mu = 1{,}68$$

O ipoteză alternativă poate fi:

$$H_1: \mu \neq 1{,}68.$$

Dar, în acest exemplu, spre deosebire de exemplul 7.1, se poate preciza şi sensul abaterii de la valoarea fixată. Se pot lua în considerare astfel şi ipotezele alternative care fixează sensul abaterii de la valoarea dată în ipoteza nulă:

$$H_1': \mu < 1{,}68 \qquad \text{sau} \qquad H_1'': \mu > 1{,}68.$$

Se spune că testul statistic caracterizat de ipotezele

$$\text{(A)} \quad \begin{cases} H_0 : \mu = 1{,}68 \text{ m} \\ H_1 : \mu \neq 1{,}68 \text{ m} \end{cases}$$

este un **test bilateral** sau **nedirecţional**, deoarece ipoteza alternativă nu fixează sensul abaterii, adică media populaţiei poate să fie mai mare sau mai mică decât ceea ce se aşteaptă.

Se spune că testul statistic caracterizat de ipotezele

$$\text{(B)} \quad \begin{cases} H_0 : \mu = 1{,}68 \text{ m} \\ H_1' : \mu < 1{,}68 \text{ m} \end{cases}$$

sau

$$\text{(C)} \quad \begin{cases} H_0 : \mu = 1{,}68 \text{ m} \\ H_1'' : \mu > 1{,}68 \text{ m} \end{cases}$$

este un **test unilateral** sau **direcţional**.

Este de remarcat faptul că, acolo unde situaţia concretă permite un test direcţional, nu are sens aplicarea a trei teste diferite, ci numai a două: un test bilateral şi acel test unilateral care este în concordanţă cu situaţia experimentală.

Revenind la exemplul 7.2, să presupunem că datele experimentale pentru un eşantion de 100 de persoane au condus la media de sondaj $\bar{x} = 1{,}70$ m.

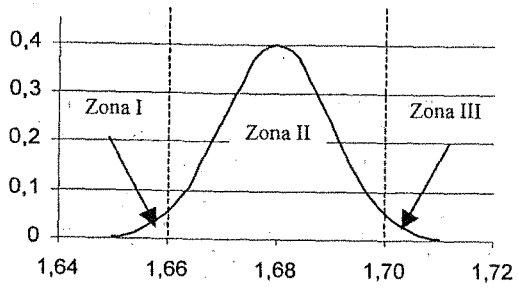


Figura 7.2. *Curba N(1,68;0,0001) şi regiunile de interes pentru un test statistic*

Indiferent de testul utilizat, considerăm că ipoteza H_0 (aceeaşi pentru toate testele A, B sau C) are loc; deci, din definiţia repartiţiei de sondaj a mediei, se poate presupune că media de sondaj provine dintr-o repartiţie normală, $N(\mu;\sigma^2/n)$, adică din repartiţia $N(1{,}68;(0{,}1)^2/100)$, altfel spus, din $N(1{,}68;0{,}0001)$. Prin urmare,

am obținut valoarea experimentală de 1,70 dintr-o repartiție cu media 1,68 m și abaterea standard de 0,01 m (= $\sqrt{0,0001}$).

Cunoașterea repartiției ne permite să calculăm probabilitatea critică a testului, în mod analog procedurii de la exemplul 7.1, unde însă calculam și repartiția. Considerăm densitatea de probabilitate normală având parametrii $\mu = 1,68$ și $\sigma = 0,01$, figura 7.2, pe care s-a poziționat și valoarea 1,70 (media eșantionului).

În funcție de valoarea observată, 1,70 m, și de depărtarea acesteia față de medie, se pot identifica trei zone sub curba normală, zone marcate pe figură:

- zona II conține valorile mai apropiate de medie decât valoarea observată;
- zona III conține valorile mai mari decât media și mai depărtate de medie decât valoarea observată;
- zona I conține valorile mai mici decât media și mai depărtate de medie decât valoarea observată.

Aceste trei regiuni permit stabilirea probabilităților critice ale testelor utilizate. Pentru fiecare dintre teste, urmărind stabilirea valorilor situate extrem în raport cu valoarea observată, se va calcula probabilitatea critică prin procedura uzuală explicată la prezentarea repartiției normale.

7.2.1. Testul bilateral (A)

Regiunea observațiilor extreme (mai depărtate de medie) în raport cu valoarea observată și care să conducă la respingerea ipotezei nule în favoarea ipotezei alternative H_1 (vezi testul A) este formată din reuniunea regiunilor I și III. Cu alte cuvinte, se iau în considerare toate valorile care sunt mai depărtate de medie decât valoarea observată, indiferent de sensul abaterii. Probabilitatea critică este, prin urmare:

$$P(\text{zona I} \cup \text{zona III}) = 0,0456.$$

Acest lucru se interpretează astfel: în condițiile ipotezei nule (media populației este de 1,68 m), există o probabilitate de 0,0456, deci de 4,56%, să obținem o medie de 1,70 m dintr-un eșantion de 100 de persoane.

Pentru $\alpha = 0,05$, se observă că p_c este mai mică decât α; prin urmare, se respinge ipoteza nulă în favoarea ipotezei alternative. Cu alte cuvinte, trebuie să acceptăm ipoteza că media populației nu mai este de 1,68 m.

Pentru $\alpha = 0,01$, se observă că p_c este mai mare decât α; prin urmare, nu se poate respinge ipoteza nulă la acest prag de semnificație.

Interpretând deciziile, în cele două cazuri de alegere a pragului α, nu trebuie să considerăm că respingerea sau tolerarea ipotezei nule este supusă arbitrariului. Între cele două situații există o diferență majoră de încredere acordată: în primul caz, încrederea este de 95% (adică 5% șansă de eroare), în al doilea caz încrederea este de 99% (adică 1% șansă de eroare). Este natural ca o încredere mai mare să fie asociată unei exigențe crescute.

7.2.2. Testul unilateral (B)

Deoarece media de sondaj este mai mare decât valoarea ipotetică, nu are sens practic să testăm H_0 contra unei ipoteze H_1, care afirmă o medie și mai mică a populației. Din punct de vedere strict formal, regiunea valorilor extreme și care duc la respingerea ipotezei nule este regiunea obținută ca reuniune a zonelor I și II. S-ar obține o probabilitate critică

$$P(\text{zona I} \cup \text{zona II}) = 0,9772.$$

Indiferent de valoarea uzuală α fixată, ipoteza nulă nu se poate respinge. Interpretarea ar fi că o medie de sondaj de 1,70 m este mai probabilă pentru o populație cu media de 1,68 m decât pentru o populație cu o medie mai mică decât 1,68 m.

7.2.3. Testul unilateral (C)

Regiunea valorilor extreme este de această dată zona III, deoarece în favoarea unei medii mai mari de 1,68 m vor fi valorile mai mari decât valoarea observată. Probabilitatea critică este

$$P(\text{zona III}) = 0,0228.$$

Dacă $\alpha = 0,05$, atunci se respinge ipoteza nulă în favoarea ipotezei alternative H_1, adică se poate considera că media populației este în realitate mai mare decât valoarea presupusă de 1,68.

La nivelul de semnificație $\alpha = 0,01$, nu se poate respinge ipoteza nulă. Reamintim că interpretarea poate fi completată prin aceea că evidența statistică a datelor de sondaj nu permite respingerea ipotezei nule.

Regiunile de respingere pentru cele trei teste se determină, de asemenea, pe baza repartiției normale (deoarece repartiția statisticii testului este normală). Dacă se fixează valoarea pragului de semnificație α, atunci se pot delimita zonele vizualizate în figura 7.3.

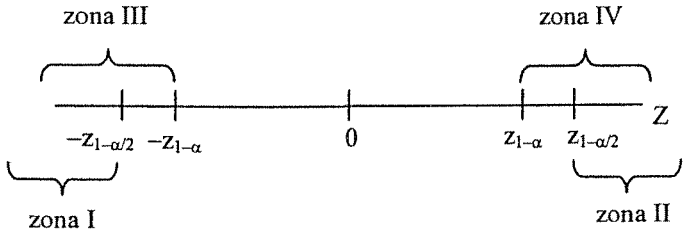


Figura 7.3. *Regiuni de respingere pentru teste bazate pe repartiția normală*

Pentru a decide în test, se efectuează mai întâi transformarea Z a mediei de sondaj:

$$Z=\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}=\frac{(\bar{x}-\mu)\sqrt{n}}{\sigma},$$

unde s-a ţinut seama de faptul că media de sondaj este repartizată $N(\mu;\sigma^2/n)$.

Având în vedere semnificaţia quantilelor repartiţiei normale standard, se poate preciza atunci că regiunile de respingere a celor trei teste sunt:

- pentru testul (A) – zona I ∪ zona II;
- pentru testul (B) – zona III;
- pentru testul (C) – zona IV.

Astfel, pentru testul (A), dacă valoarea transformată satisface $Z > z_{1-\alpha/2}$ sau $Z < -z_{1-\alpha/2}$, ceea ce se poate exprima unitar prin $|Z| > z_{1-\alpha/2}$, înseamnă că valoarea observată a mediei aparţine regiunii de respingere a ipotezei nule. În caz contrar, nu se respinge ipoteza nulă. Relaţia probabilistă este:

$$P(\text{zona I} \cup \text{zona II}) = P(\text{zona I}) + P(\text{zona II}) = \alpha/2 + \alpha/2 = \alpha,$$

adică regiunea de respingere este formată din acele valori extreme care însumează o probabilitate egală cu α. Orice valoare observată care aparţine regiunii de respingere va conduce la o regiune critică având o probabilitate mai mică decât α, deci va conduce la respingerea ipotezei nule.

Pentru testul (B), regiunea de respingere este zona caracterizată de $Z < z_{1-\alpha}$. Se ştie că $z_{1-\alpha} = -z_\alpha$ şi, prin urmare,

$$P(\text{zona III}) = P(Z < -z_{1-\alpha}) = P(Z < z_\alpha) = \alpha$$

şi se regăseşte situaţia caracteristică pentru regiunea de respingere.

În mod analog se arată valabilitatea afirmaţiei că zona IV este regiunea de respingere în testul unilateral (C).

Testul prezentat în exemplul 7.2 face parte din categoria **testelor parametrice**, deoarece ipoteza este strict legată de un parametru al populaţiei, iar statistica testului are o repartiţie cunoscută tocmai din această ipoteză.

7.3. Erori

Am accentuat în exemplele prezentate ideea că nu se pune problema obţinerii unor certitudini în domeniile analizate. Este evident atunci că se pot produce erori în procesul de decizie a unui test statistic.

Erorile pot fi de două tipuri:

- respingerea ipotezei nule, H_0, deşi H_0 este adevărată; aceste erori se numesc **erori de tip I** sau **erori de speţa întâi**.

- tolerarea (nerespingerea) ipotezei nule, H_0, deşi ea este falsă; aceste erori se numesc **erori de tip II** sau **erori de speţa a doua.**

Situaţia de eroare poate fi sintetizată în următorul tabel:

		Realitatea din populaţie	
		H_0 adevărată	H_0 falsă
Decizia în test	Respingere H_0	Eroare de tip I	Decizie justă
	Acceptare H_0	Decizie justă	Eroare de tip II

De remarcat că aceste tipuri de erori sunt similare celor cunoscute în docimologie: respingerea unui candidat, deşi acesta ar merita să fie declarat admis, sau admiterea unui candidat, deşi acesta ar merita să fie respins.

Dificultatea constă în aceea că nu există nici un mijloc practic de a şti dacă am luat o decizie corectă sau dacă am făcut o eroare şi de ce tip anume este respectiva eroare. Singurul mod raţional în care putem acţiona în asemenea situaţii este acela de a minimiza riscul de a produce erori.

Din exemplele şi noţiunile introduse până acum se observă că pragul de semnificaţie α măsoară **riscul de speţa întâi**, adică riscul de a produce o eroare de tip I. Într-adevăr, statistica testului ia, în condiţiile ipotezei nule adevărate, cu probabilitatea α, valori în regiunea de respingere. Potrivit regulii de decizie din testul statistic, această situaţie duce la respingerea ipotezei nule, deci la producerea unei erori de tip I. Ar părea necesar atunci ca nivelul de semnificaţie să fie cât mai mic. Acest fapt nu este adevărat, pentru că micşorarea riscului de speţa întâi duce la creşterea riscului de speţa a doua. Şi din acest motiv se justifică acceptarea pragurilor de 0,05 şi 0,01.

Deşi pentru o dezbatere completă trebuie consultat un tratat de statistică matematică, notăm că riscul de speţa a doua depinde de α, de reprezentativitatea eşantionului, volumul acestuia etc.

Se defineşte **puterea unui test** ca $1 - \beta$, unde β este riscul de speţa a doua, adică riscul de a tolera ipoteza nulă, deşi ea este falsă. Se poate interpreta puterea testului ca fiind capacitatea lui de a distinge o ipoteză falsă. Rezultă că pentru un test se va urmări să se minimizeze riscul de speţa întâi şi să se maximizeze puterea testului.

Deoarece în practica analizei statistice s-au construit foarte multe teste statistice, în continuare se vor prezenta doar testele utilizate frecvent şi, totodată, cele care pot constitui exemple utile pentru înţelegerea altor teste.

Pentru un novice, şi nu numai, selectarea unui test adecvat scopului urmărit devine uneori o sarcină foarte dificilă. Pentru aceştia menţionăm că un foarte util ghid de orientare în lumea testelor statistice, sub forma unui arbore de decizie, este oferit de Dirk Vorberg şi Sven Blankenberger[48].

48. D. Vorberg; S. Blankenberger, „Alegerea testelor statistice şi măsurarea", în *Psihologia socială*, Buletinul Laboratorului „Psihologia câmpului social", 4, 1999, Universitatea „Al.I. Cuza", Editura Polirom, Iaşi, pp. 121-127.

7.4. Testarea valorilor de sondaj aberante

De multe ori în cercetările experimentale se întâlneşte situaţia când anumite valori par a nu se „înscrie" în restul valorilor experimentale. Acestea sunt, de regulă, valorile extreme măsurate, ori cele maxime, ori cele minime. Apariţia unor astfel de valori, pe care le vom numi în mod obişnuit **valori aberante** sau **valori de excepţie**, se pot datora unor cauze diverse: eşantioane nereprezentative, elemente care aparţin altei populaţii statistice decât celei de interes, erori de măsurare, condiţii nefavorabile pentru prelevarea de probe etc.

Termenul de aberant are o conotaţie mai puternic negativă în vorbirea curentă decât în tematica abordată aici. În statistică, acest termen denumeşte doar valori care nu sunt în concordanţă cu celelalte date. De exemplu, pentru un elev care are la un obiect de studiu notele 10, 9, 9, 10, 10, 9, 2, 10, 9, se pare că nota 2 nu se potriveşte cu profilul lui. Această valoare poate fi socotită aberantă, fără ca prin aceasta să afirmăm că elevul nu trebuia să primească această notă, că notarea nu a fost obiectivă etc.

În privinţa valorilor aberante, sunt două aspecte care trebuie analizate:

- detectarea acestor valori;
- dacă există, acţiunea în ceea ce le priveşte: se înlătură sau se păstrează în mulţimea datelor care se analizează.

În statistică s-au dezvoltat tehnici pentru răspunsul la prima întrebare. În ceea ce priveşte al doilea aspect, doar experimentatorul poate să decidă: s-ar putea ca valoarea să fie singura de interes într-o cercetare (precum astronomul care într-o singură noapte dintr-o sută vede o cometă şi ar decide că observaţia este abeţantă).

Testele statistice dezvoltate consideră, în general, că populaţia de bază este normală, iar detectarea unei valori aberante se poate repeta după ce s-a eliminat una deja găsită.

Ca practică generală, în absenţa aplicării vreunui test, se poate lua în considerare regula care spune că, între cel puţin 10 valori, este aberantă valoarea care nu se înscrie între $\bar{x} - 4s$ şi $\bar{x} + 4s$, unde $\bar{x}$ şi s sunt calculate fără valoarea presupusă aberantă. Uneori, se ia chiar o abatere de la medie de numai $3s$. Reamintim şi criteriul 1,5×IQR, prezentat în capitolul dedicat analizei împrăştierii datelor.

Observaţie. Dată fiind utilitatea procedeului, nu numai în această secţiune, notăm un şir de valori iniţiale $x_1, x_2, \ldots, x_n$, după ordonare, prin $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$, indiferent dacă şirul este ordonat crescător sau descrescător. Notaţia este consacrată în statistică şi se va înţelege corect că este vorba despre aceleaşi valori, dar numerotate în ordinea mărimii datelor.

7.4.1. Testul Dixon

Pentru detectarea unor valori aberante într-un eşantion de volum mic, $n \leq 25$, datele de sondaj se ordonează, crescător sau descrescător, astfel încât valoarea care se testează să fie prima în şir, adică $x_{(1)}$.

Ipotezele testului Dixon sunt:

$$\begin{cases} H_0: \text{valoarea } x_{(1)} \text{ nu este aberantă;} \\ H_1: \text{valoarea } x_{(1)} \text{ este aberantă.} \end{cases}$$

Statistica testului, notată cu A, este calculată în funcţie de valoarea lui n şi este prezentată în tabelul 7.1. În acelaşi tabel sunt date şi valorile critice ale testului, notate prin $A_{n;1-\alpha}$.

Decizia în test este:

- se respinge H_0 în favoarea ipotezei H_1 dacă $A \geq A_{n;1-\alpha}$;
- nu se respinge H_0 dacă $A < A_{n;1-\alpha}$.

În cazul respingerii ipotezei nule, valoarea x(1) poate fi considerată valoare aberantă şi deci poate fi eliminată din datele procesate (dacă cercetătorul hotărăşte acest lucru). Procedura se poate relua pentru următoarea valoare, până când nu se mai respinge ipoteza nulă.

Exemplul 7.3. La selectarea unui lot de 10 elevi s-au obţinut următoarele valori reprezentând mediile la limbile străine studiate: 6,25, 6,42, 6,70, 6,85, 7,10, 7,35, 7,50, 8,33, 8,75, 9,50. Deoarece se doreşte ca lotul să fie cât mai omogen din punctul de vedere al mediilor, apare problema dacă valoarea maximă, 9,50, poate fi socotită aberantă sau nu.

Se va ordona şirul descrescător, pentru ca valoarea testată să fie prima în şir: 9,50, 8,75, 8,33, 7,50, 7,35, 7,10, 6,85, 6,70, 6,42, 6,25. Se utilizează, pentru n = 10, statistica:

$$A = \left| \frac{x_{(1)} - x_{(2)}}{x_{(1)} - x_{(n-1)}} \right| = \left| \frac{9{,}50 - 8{,}75}{9{,}50 - 6{,}42} \right| = \frac{0{,}75}{3{,}08} = 0{,}2435$$

La nivelul de semnificaţie $\alpha = 0{,}05$, linia corespunzătoare lui $n = 10$ în tabelul 7.1 arată că $A_{10;0,05} = 0{,}477$. Cum valoarea calculată este mai mică decât valoarea critică tabelată, $A = 0{,}2435 < 0{,}477$, concluzia în test este că nu se poate respinge ipoteza nulă. Prin urmare, valoarea de 9,50 nu poate fi socotită aberantă şi nu poate fi exclusă din eşantion.

Tabelul 7.1. *Testul Dixon – statistici și valori critice*

n	Valorile critice $A_{n;1-\alpha}$			Statistica A
	$\alpha = 0,10$	$\alpha = 0,05$	$\alpha = 0,01$	
3	0,886	0,941	0,988	$\frac{\lvert x_{(1)} - x_{(2)} \rvert}{\lvert x_{(1)} - x_{(n)} \rvert}$
4	0,679	0,765	0,889	
5	0,557	0,642	0,780	
6	0,482	0,560	0,698	
7	0,434	0,507	0,637	
8	0,479	0,554	0,683	$\frac{\lvert x_{(1)} - x_{(2)} \rvert}{\lvert x_{(1)} - x_{(n-1)} \rvert}$
9	0,441	0,512	0,635	
10	0,409	0,477	0,597	
11	0,517	0,576	0,679	$\frac{\lvert x_{(1)} - x_{(3)} \rvert}{\lvert x_{(1)} - x_{(n-1)} \rvert}$
12	0,490	0,546	0,642	
13	0,467	0,521	0,615	
14	0,492	0,546	0,641	$\frac{\lvert x_{(1)} - x_{(3)} \rvert}{\lvert x_{(1)} - x_{(n-2)} \rvert}$
15	0,472	0,525	0,616	
16	0,454	0,507	0,595	
17	0,438	0,490	0,577	
18	0,424	0,475	0,561	
19	0,412	0,462	0,547	
20	0,401	0,450	0,535	
21	0,391	0,440	0,524	
22	0,382	0,430	0,514	
23	0,374	0,421	0,505	
24	0,367	0,413	0,497	
25	0,360	0,406	0,489	

Valoarea minimă, 6,25, nu pare a fi neconformă cu celelalte date, deci nu s-ar impune testarea ei. Aplicarea testului Dixon, doar pentru a întări suportul intuitiv de adecvare a testului, ar trebui să diferențieze acest fapt al conformității valorii minime. Pentru testarea valorii de 6,25, se poate utiliza șirul inițial de date, care este ordonat crescător. Se va utiliza aceeași formulă pentru calculul statisticii testului (care depinde doar de volumul eșantionului). Se obține:

$$A = \left|\frac{6,25 - 6,42}{6,25 - 8,75}\right| = \left|\frac{-0,17}{-2,5}\right| = 0,068$$

Se poate vedea imediat că, oricare ar fi pragul uzual α, ipoteza nulă nu poate fi respinsă, deci valoarea minimă $x_{(1)} = 6,25$ nu este o valoare aberantă.

7.4.2. Testul Grubbs

Testul lui Grubbs se aplică eșantioanelor de volum mai mare decât 20 extrase dintr-o populație normală, cu parametrii necunoscuți. Testul se bazează pe abaterea

dintre medie și valoarea extremă testată, în unități abateri standard. Considerăm că șirul este ordonat crescător, adică $x_{(1)}$ este valoarea minimă, iar $x_{(n)}$ este valoarea maximă.

Ipotezele testului sunt identice cu cele de la testul Dixon:

$\begin{cases} H_0: \text{valoarea testată nu este aberantă;} \\ H_1: \text{valoarea testată este aberantă.} \end{cases}$

Statistica testului este:

– pentru valoarea de sondaj maximă, $x_{(n)}$

$$g = \frac{x_{(n)} - \bar{x}}{s};$$

– pentru valoarea de sondaj minimă, $x_{(1)}$

$$g = \frac{\bar{x} - x_{(1)}}{s},$$

unde $\bar{x}$ și s sunt media și, respectiv, abaterea standard de sondaj.

Ipoteza nulă se respinge dacă valoarea calculată g este mai mare decât valoarea critică tabelată a testului Grubbs. Valorile critice $g_{n;1-\alpha}$ sunt prezentate în tabelul 7.2. Analiza acestui tabel arată că pentru valori mici ale lui n se consideră că ar fi de excepție valorile observate în afara limitei de 5 abateri standard, dar domeniul admisibil se extinde spre 7 abateri standard pentru eșantioane de volum mai mare.

Tabelul 7.2. *Valorile critice pentru testul lui Grubbs,* $g_{n;1-\alpha}$

n	$\alpha = 0,05$	$\alpha = 0,01$	n	$\alpha = 0,05$	$\alpha = 0,01$
20	2,557	2,884	85	3,151	3,543
25	2,683	3,009	90	3,171	3,563
30	2,745	3,103	95	3,189	3,582
35	2,811	3,178	100	3,207	3,600
40	2,866	3,240	105	3,224	3,617
45	2,914	3,292	110	3,239	3,632
50	2,956	3,336	115	3,254	3,647
55	2,992	3,376	120	3,267	3,662
60	3,025	3,411	125	3,281	3,675
65	3,055	3,442	130	3,294	3,688
70	3,082	3,471	135	3,306	3,700
75	3,107	3,496	140	3,318	3,712
80	3,130	3,521	145	3,328	3,723

Exemplul 7.4. Să presupunem că datele de sondaj pentru exemplul 7.2, care au condus la o medie de sondaj de 1,70 m, sunt caracterizate şi de o abatere standard de sondaj de $s = 0{,}105$ m. Dacă în eşantionul ales există o persoană cu înălţimea de 2,20 m, se pune problema dacă această valoare este sau nu o valoare de excepţie care ar trebui să fie, eventual, eliminată din analiză. Reamintim că eşantionul are un volum n = 100.

Deoarece valoarea testată este cea maximă, se utilizează statistica:

$$g = \frac{x_{(100)} - \bar{x}}{s} = \frac{2{,}20 - 1{,}70}{0{,}105} = \frac{0{,}50}{0{,}105} = 4{,}762$$

şi, cum din tabelul 7.2 se găseşte $g_{100;1-0,01} = g_{100;0,99} = 3{,}600$, rezultă că la nivelul de semnificaţie de 1% se poate respinge ipoteza nulă. Deci valoarea 2,20 m poate fi considerată de excepţie în cadrul eşantionului ales.

7.5. Teste de concordanţă

Testele prezentate în această secţiune se referă la potrivirea, concordanţa dintre valorile calculate în eşantion (statisticile de sondaj) şi valorile parametrilor respectivi din populaţia statistică (valori cunoscute sau presupuse). Cu alte cuvinte, problema poate fi formulată astfel: *cât de mult poate să se abată o valoare calculată dintr-un eşantion de la valoarea presupusă în cazul întregii populaţii pentru a putea considera că are loc o nepotrivire între cele două valori.*

Testele de concordanţă mai sunt cunoscute şi sub denumirea de teste de ajustare, care exprimă potrivirea unei distribuţii empirice la o distribuţie aşteptată sau teoretică.

Deşi, formulată astfel, problema pare a se referi în primul rând la eşantion, şi nu la populaţia de bază, punctul de vedere corect este explicat în următoarele afirmaţii:

- există o populaţie statistică de interes, fie ea P_1;
- pentru orice eşantion se poate considera o populaţie de bază din care eşantionul este extras; fie P_2 această populaţie;
- problema este dacă se poate considera că P_2 concordă cu P_1.

Se observă că testarea se va efectua pentru ipoteze privind populaţii, se va utiliza informaţia dintr-un eşantion, deci rămânem în domeniul inferenţei statistice.

Testele prezentate sunt clasificate în teste pentru variabile continue şi teste pentru variabile discrete (nominale sau ordinale). Primele sunt, de regulă, teste parametrice, celelalte fiind neparametrice. Este inadecvat să se utilizeze teste proiectate pentru date continue la date discrete. Situaţia inversă este posibilă prin aceea că, pe de o parte, se pot discretiza variabilele continue, iar, pe de altă parte, testele neparametrice sunt mai puţin restrictive la condiţiile de aplicare sau, de multe ori, restricţiile impuse de un test parametric nu sunt îndeplinite. Reversul

medaliei este acela că testele neparametrice sunt mai puţin sensibile la diferenţe decât testele parametrice.

Ipoteza nulă va afirma, în general, că populaţiile P_1 şi P_2 concordă. Respingerea ipotezei nule poate avea, în practică, două consecinţe:

- se va considera că eşantionul nu este reprezentativ pentru populaţia de interes, populaţie considerată stabilă; se va preleva un alt eşantion;
- se va considera că populaţia P_1 s-a schimbat; noua populaţie de referinţă este P_2.

Opţiunea pentru una dintre cele două afirmaţii aparţine practicianului din domeniul de aplicaţie, fiind, de cele mai multe ori, o alegere ghidată de intuiţie, de experienţă etc.

7.5.1. Testul erorii standard a mediei

Fie P_1 populaţia statistică de interes, caracterizată de media μ_0 (cunoscută sau presupusă) şi de abaterea standard σ (cunoscută). Prin prelucrarea datelor unui eşantion de volum n, se obţin media de sondaj $\bar{x}$ şi abaterea standard de sondaj s. Întrebarea la care se caută un răspuns este dacă valorile tipice de sondaj susţin ipoteza că eşantionul face parte din populaţia P_1, accentul fiind pus pe media populaţiei.

În testul erorii standard a mediei, se presupune că sunt îndeplinite condiţiile care asigură mediei de sondaj o repartiţie normală sau aproape normală: caracteristica studiată este repartizată normal sau eşantionul este mare ($n \geq 30$). Se ştie că, în aceste condiţii, media de sondaj urmează o repartiţie normală $N(\mu, \sigma^2/n)$, unde μ este media populaţiei (notată în introducerea secţiunii cu P_2) din care provine eşantionul. Se observă că pentru P_2 se presupune aceeaşi abatere standard σ. Rezultă că variabila transformată

$$Z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu}{\sigma}\sqrt{n}$$

este repartizată normal standard şi poate fi utilizată pentru calcularea probabilităţilor necesare.

Ipotezele sunt:

- pentru testul bilateral:

$$(A)\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases};$$

- pentru testele unilaterale:

$$(B)\begin{cases} H_0 : \mu = \mu_0 \\ H_1' : \mu < \mu_1 \end{cases} \text{ sau } (C)\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{cases},$$

unde s-au păstrat notaţiile de la exemplul 7.2.

În condițiile ipotezei nule, $\mu = \mu_0$, rezultă că transformata Z a mediei de sondaj devine:

$$Z = \frac{\bar{x} - \mu_0}{\sigma}\sqrt{n},$$

în care toate valorile sunt cunoscute și, prin urmare, valoarea Z obținută poate fi localizată pe curba densității de probabilitate normală standard.

Deoarece procedura de decizie într-un asemenea test este explicată pe larg în exemplul 7.2, prezentăm doar regula de decizie, la pragul de semnificație fixat α:

Testul bilateral (A):

- se respinge ipoteza nulă H_0 în favoarea ipotezei alternative H_1 dacă valoarea Z calculată este mai mare, în valoare absolută, decât quantila $z_{1-\alpha/2}$ a repartiției normale standard:

$$|Z| \geq z_{1-\alpha/2};$$

- nu se respinge ipoteza nulă în caz contrar:

$$|Z| < z_{1-\alpha/2}.$$

Testul unilateral (B):

- se respinge ipoteza nulă H_0 în favoarea ipotezei alternative H_1' dacă valoarea Z calculată este mai mică decât quantila $(-z_{1-\alpha})$ a repartiției normale standard:

$$Z \leq -z_{1-\alpha};$$

- nu se respinge ipoteza nulă în caz contrar:

$$Z > -z_{1-\alpha}.$$

Testul unilateral (C):

- se respinge ipoteza nulă H_0 în favoarea ipotezei alternative H_1'' dacă valoarea Z calculată este mai mare decât quantila $z_{1-\alpha}$ a repartiției normale standard:

$$Z \geq z_{1-\alpha};$$

- nu se respinge ipoteza nulă în caz contrar:

$$Z < z_{1-\alpha}.$$

Concluzii. Pentru a aplica acest test, este necesar să se cunoască σ; prin urmare, situația practică de referință este aceea în care se cercetează dacă o populație statistică, de variabilitate constantă, și-a menținut sau nu valoarea medie.

Deoarece, în general, nu se poate ști cu siguranță că repartiția caracteristicii studiate este riguros normală, acest test se utilizează pentru eșantioane mari, beneficiind astfel de consecințele teoremei de limită centrală.

Dacă se calculează, la nivelul α considerat, intervalul de încredere a mediei populației, decizia în test se poate lua și în funcție de faptul că valoarea μ_0 aparține sau nu intervalului de încredere: se respinge H_0 dacă μ_0 nu aparține intervalului estimat.

Acest test este utilizat și cu denumirea de testul Z de concordanță, datorită utilizării unei statistici repartizate normal standard.

7.5.2. Testul de concordanță Student (t)

Atunci când nu se cunoaște abaterea standard a populației, σ, testul anterior nu mai poate fi aplicat. În această situație se va utiliza estimația *s*, abaterea standard de sondaj, în locul lui σ, iar repartiția auxiliară va fi repartiția Student.

Pentru caracteristica studiată se presupune însă o repartiție normală (cu parametri necunoscuți) sau apropiată de una normală.

Ipotezele testului sunt aceleași cu seturile de ipoteze anterioare (A), (B), (C).

Statistica testului este similară statisticii din testul erorii standard a mediei, cu excepția faptului că, în loc de σ, se utilizează estimația *s*:

$$t = \frac{\bar{x} - \mu_0}{s}\sqrt{n}.$$

Dacă ipoteza nulă, H_0: $\mu = \mu_0$, este adevărată, atunci variabila t urmează o repartiție Student cu $\nu = n - 1$ grade de libertate. Prin urmare, decizia în test se ia după mecanismul general al unui test pentru care se cunoaște repartiția statisticii testului (vezi exemplul 7.2). Pentru aplicațiile practice, precizăm că, în testul t, se consultă tabelul statistic cu quantilele repartiției Student.

Testul bilateral (A):

- se respinge H_0 în favoarea ipotezei alternative H_1: $\mu \neq \mu_0$, dacă valoarea calculată, t, este mai mare în valoare absolută decât quantila de ordin $1 - \alpha/2$ a repartiției Student corespunzătoare:

$$|t| \geq t_{1-\alpha/2;\nu};$$

- nu se respinge H_0 în cazul contrar:

$$|t| < t_{1-\alpha/2;\nu}.$$

Testul unilateral (B):

- se respinge ipoteza nulă în favoarea ipotezei alternative, H_1': $\mu < \mu_0$, dacă valoarea calculată este mai mică decât quantila de ordin α (egală cu valoarea simetrică a quantilei de ordin $1 - \alpha$) a distribuției Student corespunzătoare:

$$t \leq -t_{1-\alpha;\nu};$$

- nu se respinge H_0 în cazul contrar:

$$t > -t_{1-\alpha;\nu}.$$

Testul unilateral (C):

- se respinge ipoteza nulă în favoarea ipotezei alternative, $H_1': \mu > \mu_0$, dacă valoarea calculată este mai mare decât quantila de ordin $1-\alpha$ a distribuției Student corespunzătoare:

$$t \geq t_{1-\alpha;\nu};$$

- nu se respinge H_0 în cazul contrar:

$$t < t_{1-\alpha;\nu}.$$

Concluzii. Acest test se va utiliza atunci când nu este cunoscută dispersia populației, ceea ce îl deosebește de testul erorii standard a mediei (când dispersia populației este cunoscută).

Testul t poate fi utilizat și pentru eșantioane mici, dar trebuie cunoscut faptul că repartiția caracteristicii de interes este normală (sau foarte apropiată de repartiția normală).

7.5.3. Testul erorii standard a unei proporții

Acest test este aplicabil în cazul în care se cercetează o variabilă dihotomică și se dorește răspunsul la întrebarea: *s-a schimbat sau nu proporția elementelor de interes în populație*?

Se știe că asemenea caracteristici sunt modelate prin variabile repartizate binomial, proporția cazurilor notate „succes" fiind asimilată parametrului p din distribuția binomială, Bi(n;p). Din această cauză testul este cunoscut și ca test de concordanță pentru parametrul p al repartiției binomiale.

Să luăm o populație în care proporția cazurilor „succes" (vezi convenția făcută la prezentarea repartiției binomiale) este p_0. Dispunem de un eșantion de volum n în care frecvența cazurilor „succes" este r. Mai notăm cu p, în spiritul raționamentului prezentat la testele anterioare, proporția cazurilor „succes" în populația „adevărată" de unde este extras eșantionul.

Ipotezele care traduc problema inițială sunt, pentru testul bilateral,

$$(A)\begin{cases} H_0 : p = p_0 \\ H_1 : p \neq p_0 \end{cases},$$

iar pentru testele unilaterale,

$$(B)\begin{cases} H_0 : p = p_0 \\ H_1' : p < p_0 \end{cases}; \quad (C)\begin{cases} H_0 : p = p_0 \\ H_1'' : p > p_0 \end{cases}.$$

Cu alte cuvinte, ipoteza H_0 afirmă conservarea proporției cazurilor „succes" (concordanța între eșantion și populația presupusă), ipoteza H_1 afirmă că proporția s-a modificat (fără a preciza sensul modificării), iar ipotezele alternative din testele unilaterale precizează sensul transformării populației.

Situația I. În cazul în care n este suficient de mare (a se vedea discuția de la repartiția binomială), se poate recurge la aproximarea normală a repartiției binomiale. În acest caz, dacă H_0 este adevărată, atunci variabila

$$Z = \frac{r - np_0}{\sqrt{np_0(1-p_0)}} = \frac{\frac{r}{n} - p_0}{\sqrt{p_0(1-p_0)}}\sqrt{n}$$

are aproximativ o repartiție normală standard. Se va observa că expresia lui Z este construită tot pe baza diferenței dintre ceea ce este observat și ceea ce este așteptat: r/n reprezintă frecvența relativă a cazurilor „succes" în eșantion (deci o estimare a lui p_0), după cum np_0 reprezintă frecvența așteptată a cazurilor „succes" în eșantion.

Decizia în test este similară unui test bazat pe N(0;1). Se respinge ipoteza nulă H_0 în favoarea ipotezelor alternative, dacă sunt îndeplinite, corespunzător seturilor de ipoteze (A) – (C), condițiile:

$$(A)\ |Z| \geq z_{1-\alpha/2};\ (B)\ Z \leq -z_{1-\alpha};\ (C)\ Z \geq z_{1-\alpha}.$$

Situația II. În cazul în care n este relativ mic, se poate efectua un calcul similar celui prezentat în testul neparametric din exemplul 7.1. De exemplu, pentru testul (C), evenimentele situate extrem în raport cu cel realizat și care conduc la respingerea ipotezei nule sunt aparițiile a mai mult de r „succese". Probabilitatea evenimentului compus este, prin urmare, sub H_0,

$$P_{H_0} = \sum_{i=r}^{n} C_n^i p_0^i (1-p_0)^{n-i},$$

în care se însumează probabilitățile de apariție a evenimentelor caracterizate de un număr de „succese" mai mare sau egal cu r. Se obține astfel probabilitatea critică a testului, adică probabilitatea de a obține un număr de „succese" mai mare sau egal cu r.

Potrivit discuției privind probabilitățile critice, se va respinge ipoteza nulă dacă probabilitatea critică este mai mică decât nivelul de semnificație α fixat, $P_{H_0} \leq \alpha$.

În mod analog se pot calcula probabilitățile critice și pentru testele (A) și (B). Este necesar ca n să fie suficient de mic pentru a se efectua fără erori calculul combinărilor din formulă.

7.5.4. Testul Kolmogorov-Smirnov (K-S)

În cazul datelor discrete (cu excepția celor dihotomice), nu se poate aplica nici unul dintre testele parametrice prezentate anterior. Studiul concordanței unei distribuții de date discrete este, pe de altă parte, deosebit de studiul concordanței pentru date continue. În primul rând, valorile tipice uzuale pentru date continue, media și

abaterea standard, nu mai sunt definite pentru datele discrete. În al doilea rând, o scală de măsură discretă conține relativ puține categorii, în așa fel încât ar prezenta interes o comparare globală a situației descrise de un eșantion cu o situație așteptată.

Pentru datele ordinale este posibilă o asemenea rezolvare utilizând testul Kolmogorov-Smirnov. Acest test neparametric nu este dezvoltat special pentru date ordinale. El poate fi aplicat și pentru studiul concordanței datelor continue, caz în care necesită eșantioane de volum mai mic decât pentru date ordinale.

Testul se bazează pe compararea funcției de repartiție de sondaj (reprezentată de frecvențele relative cumulate) cu funcția de repartiție teoretică (așteptată). Rămânând în domeniul datelor ordinale, situația este explicată în tabelul 7.3:

Tabelul 7.3. *Explicații la testul Kolmogorov-Smirnov*

Categoria	Frecvența observată		Frecvența așteptată	Diferența absolută
	absolută	cumulată		
c_1	f_1	$F^*(c_1)$	$F(c_1)$	$\lvert F^*(c_1) - F(c_1) \rvert$
c_2	f_2	$F^*(c_2)$	$F(c_2)$	$\lvert F^*(c_2) - F(c_2) \rvert$
...	...	...	...	...
c_k	f_k	$F^*(c_k)$	$F(c_k)$	$\lvert F^*(c_k) - F(c_k) \rvert$

Se observă că se compară, prin diferențele calculate, frecvențele relative cumulate. Se realizează astfel o comparare individuală, după fiecare rang al scalei, dar și o comparare globală, prin faptul că frecvențele sunt cumulate. Statistica testului este valoarea maximă a diferenței absolute:

$$D = \max_i \left| F^*(c_i) - F(c_i) \right|.$$

Ipotezele testului K-S sunt de tip bilateral:

$\begin{cases} H_0: \text{distribuția populației coincide cu distribuția așteptată;} \\ H_1: \text{distribuția populației este diferită de distribuția așteptată.} \end{cases}$

Regula de decizie în test se bazează pe compararea valorii D calculate cu valoarea critică ce corespunde în testul K-S. Valorile critice sunt prezentate[49], pentru două niveluri de semnificație uzuale, în tabelul 7.4.

Se respinge ipoteza nulă dacă valoarea calculată D este mai mare sau egală cu valoarea tabelată D*, corespunzătoare volumului n al eșantionului și nivelului de semnificație α fixat.

49. Aparatul matematic implicat este suficient de complicat pentru a depăși nivelul asumat la prezentarea acestei cărți.

Tabelul 7.4. *Valorile critice D* pentru testul K-S (un eșantion)*

n	$\alpha = 0{,}05$	$\alpha = 0{,}01$	n	$\alpha = 0{,}05$	$\alpha = 0{,}01$
3	0,708	0,828	14	0,349	0,418
4	0,624	0,733	15	0,338	0,404
5	0,565	0,669	16	0,328	0,392
6	0,521	0,618	17	0,318	0,381
7	0,486	0,577	18	0,309	0,371
8	0,457	0,543	19	0,301	0,363
9	0,432	0,514	20	0,294	0,356
10	0,410	0,490	25	0,27	0,32
11	0,391	0,468	30	0,24	0,29
12	0,375	0,450	35	0,23	0,27
13	0,361	0,433	> 35	$1{,}36/\sqrt{n}$	$1{,}63/\sqrt{n}$

Testul K-S este neparametric (nu se referă la un parametru anume al repartiției) și este independent de distribuție (nu se fac presupuneri speciale asupra repartiției populației). În cazul în care nu se respinge ipoteza H_0, se poate accepta că repartițiile sunt identice, ceea ce presupune egalitatea unor parametri sau a formei repartiției etc.

Deoarece este considerat un singur eșantion, testul prezentat se mai numește testul K-S pentru un eșantion (de unde și titlul tabelului 7.4). Există și testul K-S pentru două eșantioane, prezentat la testele de comparare.

Aplicarea testului K-S în cazul unei variabile continue impune gruparea prealabilă a datelor de sondaj într-un număr de intervale (vezi prezentarea datelor continue). Apoi se procedează în mod similar procedurii pentru date ordinale.

7.5.5. Testul de concordanță χ^2

Pentru a verifica apropierea de o repartiție cunoscută sau presupusă a unei variabile nominale cu mai mult de două categorii, se utilizează cu predilecție testul χ^2, un test de aceeași factură cu testul Kolmogorov-Smirnov. Testul este neparametric, independent de distribuție și privește repartiția în totalitatea ei.

În aplicarea testului pentru o caracteristică nominală este utilizată noțiunea de valoare așteptată (nu numai aici, de altfel), care trebuie gândită în sensul comun în care un gospodar, care își pregătește pentru toamnă două butoaie mari pentru vin și un butoiaș pentru țuică, se așteaptă să obțină mai mult vin decât țuică. În test, valorile așteptate sunt valori cunoscute din studii anterioare sau valori presupuse (așteptate în sensul explicat anterior).

Fie o variabilă nominală cu un număr r de categorii. Notăm cu E_i frecvența așteptată în categoria i, $i = 1, \ldots, r$. Notăm cu O_i, $i = 1, \ldots, r$, frecvențele observate într-un eșantion pentru categoriile nominale.

Se pune problema să se compare distribuția observată cu distribuția așteptată. Ipotezele testului sunt de tip bilateral, similare celor de la testul K-S:

$\begin{cases} H_0: \text{distribuția populației coincide cu distribuția așteptată;} \\ H_1: \text{distribuția populației este diferită de distribuția așteptată.} \end{cases}$

Prin distribuția populației se va înțelege corect distribuția populației pentru care eșantionul este reprezentativ.

Statistica testului este definită prin:

$$\chi^2 = \sum_{i=1}^{r} \frac{(O_i - E_i)^2}{E_i},$$

care, în condițiile ipotezei H_0, este repartizată χ^2 cu $\nu = r - 1$ grade de libertate. Se explică astfel și de ce statistica este denumită χ^2.

Regula de decizie. Ipoteza nulă se respinge, la nivelul de semnificație α fixat, dacă valoarea calculată este mai mare sau egală cu valoarea critică tabelată $\chi^2_{1-\alpha;r-1}$.

În aplicarea testului se impun anumite condiții:

- toate frecvențele așteptate trebuie să fie diferite de zero, $E_i \neq 0, i = 1, \ldots, r$,
- cel mult 20% dintre categorii pot să aibă frecvențe observate mai mici de 5.

Testul de concordanță χ^2 poate fi utilizat, la fel ca și testul K-S, și pentru date ordinale, și pentru cele continue. În acest caz, se recurge la gruparea inițială a datelor într-un număr de clase, potrivit procedurii obișnuite. Dacă a doua condiție nu este îndeplinită, se poate proceda la regruparea datelor în clase mai mari (această regrupare nu se poate utiliza, în general, pentru situația datelor nominale).

Regulă de calcul. Distribuția așteptată poate fi cunoscută sub formă de proporții, situație frecventă pentru datele continue, când se verifică ipoteza concordanței cu o repartiție teoretică. Pentru a reveni la frecvențele absolute așteptate se utilizează

$$E_i = n \times p_i,\ i = 1, \ldots, r,$$

unde n este volumul eșantionului, iar p_i este proporția (sau probabilitatea) așteptată pentru clasa i.

Dacă în estimarea proporțiilor așteptate se utilizează o repartiție teoretică, pentru care s-au estimat din eșantion k parametri, atunci repartiția statisticii testului este repartiția χ^2 cu $\nu = r - (k + 1)$ grade de libertate. Acest număr de grade de libertate este utilizat în găsirea valorilor critice $\chi^2_{1-\alpha;r-(k+1)}$. Pentru k = 0, se aplică formula indicată inițial.

De exemplu, dacă dintr-un eșantion se estimează media și abaterea standard pentru estimarea parametrilor repartiției normale, atunci $k = 2$ și se va utiliza, ca repartiție auxiliară, χ^2 cu $\nu = r - (2+1) = r - 3$ grade de libertate, r fiind numărul de clase de grupare.

Observație. Pentru o discuție mai amănunțită a aplicării testelor χ^2 în domeniul cercetărilor psihologice, se va consulta secțiunea „Utilizarea testelor χ^2 în psihologie" din finalul acestui capitol.

7.6. Teste de comparare

Categoriile de teste prezentate drept teste de comparare se bazează pe compararea datelor de sondaj care aparțin la două eșantioane. După cum s-a văzut în întreaga dezbatere cu privire la eșantioane și sondaje, șansa de a fi obținute două eșantioane identice este extrem de redusă. Din această cauză, problema comparării eșantioanelor, luată în sensul strict al cuvântului, pare lipsită de importanță, deoarece întotdeauna vor exista diferențe între diverse eșantioane.

Vom înțelege însă corect problematica privind compararea eșantioanelor dacă o vom înscrie în inferența statistică. Fie astfel două eșantioane reprezentative, pentru două populații P_1 și P_2. Prin compararea eșantioanelor se dorește, de fapt, compararea celor două populații. Dificultatea procedurii constă în aceea că diferențele dintre cele două eșantioane se pot datora:

- diferențelor dintre populații și/sau
- diferențelor de sondaj dintre eșantioane.

Reversul medaliei este că similaritatea între eșantioane se poate datora fie egalității populațiilor, fie întâmplării.

Dacă diferențele care se detectează din datele empirice sunt considerate prea mari, atunci se va conchide că faptul nu se poate datora operației de sondaj (deci întâmplării), ci reflectă o diferențiere între populații. Dacă diferențele empirice sunt considerate mici, atunci nu se poate susține, pe baza datelor de sondaj analizate, diferența dintre cele două populații.

7.6.1. Situații de aplicare a testelor de comparare

Cele două populații, P_1 și P_2, care se compară, pot fi identificate în mai multe situații experimentale. Concluziile sunt diferite după situația întâlnită, deci este importantă identificarea exactă a cazului de aplicare a comparației.

Cazul I. Populații diferite sunt comparate pentru a decide dacă o anumită caracteristică este repartizată la fel în cele două populații. Procedura are sens dacă între populații există, pe cât posibil, cât mai multe atribute comune, ceea ce le conferă caracterul de subpopulații ale unei populații mai vaste. Identificarea celor două populații se face pe baza unei variabile de clasificare, în general nominală.

În cazul în care concluzia testului este că populațiile sunt identice, faptul se interpretează prin aceea că variabila testată nu depinde (sau nu este asociată) de variabila de clasificare. În cazul în care se decide existența unei deosebiri între

populații, faptul se interpretează ca fiind asociere între variabila testată și variabila de clasificare; altfel spus, variabila testată are valori diferite pentru modalități diferite ale variabilei de clasificare.

Exemplul 7.5. Dacă se compară două eșantioane de elevi, unul din mediul urban și unul din mediul rural, după valorile coeficientului de inteligență, concluziile vor privi *populația rurală de elevi* și *populația urbană de elevi.* Sunt posibile două concluzii (evident, probabiliste):

- cele două populații sunt identice (sau nu se poate susține existența vreunei deosebiri); în acest caz, interpretăm că respectivul coeficient de inteligență nu depinde de variabila mediu de reședință (urban/rural) sau că între cele două variabile nu există asociere;
- cele două populații sunt diferite; în acest caz, interpretăm că indicele de inteligență depinde de mediul de reședință, adică variabilele mediu de reședință și coeficient de inteligență sunt asociate.

Cazul II. Aceeași populație este evaluată la momente diferite în timp sau în situații, în condiții diferite. Evaluarea se efectuează prin eșantioane diferite. În general, un eșantion este supus unui experiment, tratament etc. și este denumit în mod obișnuit **grup experimental**. Celălalt eșantion nu este supus experimentului și joacă rolul de **grup de control**. Compararea efectuată între cele două eșantioane are drept scop stabilirea efectului experimentului efectuat; se stabilește deci dacă experimentul produce o modificare semnificativă în repartiția grupului experimental, încât acesta poate fi considerat ca aparținând unei populații noi, diferită de populația inițială.

O noțiune importantă în compararea eșantioanelor, cu implicații în tratarea statistică, este aceea de independență a eșantioanelor. Se spune că două eșantioane sunt **independente** sau **necorelate** dacă selectarea elementelor unui eșantion nu este influențată și nici nu influențează selectarea elementelor celuilalt eșantion. În caz contrar, eșantioanele sunt **dependente**, **corelate** sau **perechi**. De exemplu, selectarea a două grupuri de elevi care să fie similare după ocupația părinților (dacă primul grup conține un elev cu tatăl cadru militar, atunci și al doilea grup trebuie să conțină un elev cu tatăl cadru militar etc.) produce grupuri corelate. Cu alte cuvinte, selectarea elementelor unui grup determină caracteristicile elementelor unui grup pereche.

Cazul frecvent de apariție a unor grupuri corelate este acela când dintr-o populație se selectează un eșantion. Acest eșantion este măsurat de două ori, între cele două măsurători fiind supus unui experiment (sau unui test nestatistic). Scorurile pre-test și post-test sunt comparate pentru a decide asupra efectului experimentului. Compararea celor două seturi de date va fi efectuată prin tehnici statistice dedicate datelor corelate.

Drept consecință, testele statistice de comparare sunt diferențiate atât după natura variabilelor (continue, discrete), cât și după natura eșantioanelor (independente, dependente).

Datorită logicii expunerii, începem prin a prezenta un test de comparare a dispersiilor, testul F (Fisher-Snedecor).

7.6.2. Testul F

Compararea mediilor populațiilor normale ia în considerare împrăștierea datelor în cele două populații. Este important atunci să se cunoască dacă dispersiile celor două populații pot fi considerate egale sau nu. Acest fapt se decide utilizând testul F, bazat pe repartiția teoretică F (Fisher-Snedecor).

Situația poate fi identificată prin:

- două populații, caracterizate de variabilele X_1 și, respectiv, X_2,;
- variabilele sunt repartizate normal, $X_1 \sim N(\mu_1;\sigma_1^2)$, $X_2 \sim N(\mu_2;\sigma_2^2)$;
- din două eșantioane, câte unul din fiecare populație, dispunem de estimațiile s_1^2 și s_2^2 ale dispersiilor populațiilor; eșantioanele au volume n_1 și, respectiv, n_2.

Ipotezele testului F sunt atât de tip bilateral, cât și de tip unilateral:

Testul bilateral:

$$\text{(A)} \quad \begin{cases} H_0 : \sigma_1^2 = \sigma_2^2 \\ H_1 : \sigma_1^2 \neq \sigma_2^2 \end{cases}$$

Teste unilaterale:

$$\text{(B)} \quad \begin{cases} H_0 : \sigma_1^2 = \sigma_2^2 \\ H_1' : \sigma_1^2 < \sigma_2^2 \end{cases}; \quad \text{(C)} \quad \begin{cases} H_0 : \sigma_1^2 = \sigma_2^2 \\ H_1'' : \sigma_1^2 > \sigma_2^2 \end{cases}.$$

Când ipoteza nulă este adevărată, atunci statistica

$$F^* = \frac{s_1^2}{s_2^2}$$

este repartizată F cu $\nu_1 = n_1 - 1$ și $\nu_2 = n_2 - 1$ grade de libertate, astfel încât se pot utiliza valorile tabelate pentru $F(\nu_1;\nu_2)$ pentru determinarea probabilităților critice.

Pentru simplificarea deciziei în test, în practică se utilizează o statistică ușor modificată prin luarea în considerare ca primă populație, P_1, a populației pentru care dispersia de sondaj este mai mare:

$$F = \frac{\max(s_1^2, s_2^2)}{\min(s_1^2, s_2^2)},$$

în așa fel încât sunt utilizabile doar testele (A) și (C). În acest caz, se notează cu ν_{max} numărul gradelor de libertate pentru numărător și cu ν_{min} numărul gradelor de libertate pentru numitor.

Decizia, la nivelul de semnificaţie α, pentru testul bilateral (A):

- se respinge ipoteza nulă H_0 în favoarea ipotezei alternative H_1 dacă:

$$F > F_{1-\alpha/2;\nu_{max};\nu_{min}} \text{ sau } F < F_{\alpha/2;\nu_{max};\nu_{min}};$$

- nu se respinge ipoteza nulă, în caz contrar.

Decizia, la nivelul de semnificaţie α, pentru testul unilateral (C):

- se respinge ipoteza nulă H_0 în favoarea ipotezei alternative H_1' dacă:

$$F > F_{1-\alpha;\nu_{max};\nu_{min}};$$

- nu se respinge ipoteza nulă, în caz contrar.

Exemplul 7.6. Să presupunem că pentru un eşantion de 20 de elevi din mediul rural s-a obţinut o dispersie de sondaj a înălţimii egală cu 0,0064 (adică o abatere standard de 8 cm), iar pentru un eşantion de 30 de elevi din mediul urban s-a obţinut o dispersie de sondaj egală cu 0,01 (adică o abatere standard de 10 cm). Trebuie să decidem, la nivelul de semnificaţie $\alpha = 0{,}05$, dacă populaţiile de elevi au aceeaşi dispersie a înălţimii sau nu.

Cu notaţiile testului avem:

$$s_1^2 = 0{,}01;\ s_2^2 = 0{,}0064;\ F = \frac{s_1^2}{s_2^2} = 1.5625;\ \nu_{max} = 29;\ \nu_{min} = 19.$$

De remarcat că s-a ales drept primă populaţie cea cu dispersia de sondaj maximă, adică populaţia urbană.

La nivelul α impus, în tabelul repartiţiei F se determină valorile critice $F_{0,975;29;19} = 2{,}402$ şi $F_{0{,}025;29;19} = 0{,}44$. Potrivit procedurii de decizie, se observă că valoarea F calculată se situează între valorile critice; prin urmare, nu se poate respinge ipoteza nulă în testul bilateral (A). Rezultă că diferenţa observată dintre dispersiile de sondaj se datorează, cu o probabilitate mai mare, sondajului, şi nu unei diferenţe între dispersiile populaţiilor. Datele de sondaj nu susţin ipoteza unor dispersii diferite.

Dacă testul bilateral nu respinge ipoteza nulă, nu se mai verifică testul unilateral (C). Se poate tolera ipoteza că împrăştierea înălţimilor nu diferă după mediul de proveniență (rural sau urban) al elevilor.

***

Următoarele două teste, bazate pe repartiţia auxiliară Student, permit verificarea ipotezei de egalitate a mediilor pentru două populaţii, în cazul în care nu se cunosc dispersiile populaţiilor. Testele sunt diferenţiate prin egalitatea sau inegalitatea dispersiilor, necunoscute, ale celor două populaţii, fapt decis printr-un test F.

7.6.3. Testul Student pentru eşantioane independente, abateri standard egale

În cazul în care nu se poate respinge ipoteza egalităţii dispersiilor celor două populaţii, datorită aplicării unui test F sau datorită cunoaşterii din altă sursă a egalităţii dispersiilor, situaţia poate fi schematizată prin:

- populaţia P_1 cu media μ_1 şi dispersia σ^2;
- populaţia P_2 cu media μ_2 şi dispersia σ^2 (aceeaşi ca la P_1);
- o variabilă X repartizată normal în cele două populaţii, cu parametrii corespunzători menţionaţi;
- un eşantion din P_1, de volum n_1, cu media $\bar{x}_1$ şi dispersia s_1^2;
- un eşantion din P_2, de volum n_2, cu media $\bar{x}_2$ şi dispersia s_2^2.

Se presupune că eşantioanele sunt independente.

În aceste condiţii, cele de normalitate fiind importante, variabila:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} \times \sqrt{\frac{n_1+n_2}{n_1 n_2}}}$$

are o repartiţie Student cu $\nu = n_1 + n_2 - 2$ grade de libertate.

Această statistică permite compararea cu o diferenţă prestabilită a mediilor populaţiilor (deci cu o valoare fixată pentru $\mu_1 - \mu_2$). În cele mai multe aplicaţii reale se ia însă o diferenţă egală cu zero, adică egalitatea mediilor. Ipotezele avute în vedere în acest caz sunt:

$$\text{(A)} \begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases}, \text{ pentru testul bilateral,}$$

$$\text{(B)} \begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1' : \mu_1 < \mu_2 \end{cases} \text{ sau (C)} \begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1'' : \mu_1 > \mu_2 \end{cases}, \text{ pentru testul unilateral.}$$

În cazul în care ipoteza H_0 este adevărată, adică $\mu_1 - \mu_2 = 0$, expresia statisticii testului se simplifică, devenind:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} \times \sqrt{\frac{n_1+n_2}{n_1 n_2}}}.$$

Numitorul mai este notat $s_{\bar{x}_1 - \bar{x}_2}$ şi se înţelege ca fiind eroarea standard a diferenţei mediilor de sondaj.

Regula de decizie este aceeaşi cu a unui test t:

- se respinge H_0 în favoarea ipotezei H_1 (în testul bilateral A), dacă:

$$|t| \geq t_{1-\alpha/2;\nu};$$

- se respinge H_0 în favoarea ipotezei H_1' (în testul unilateral B), dacă:

$$t < -t_{1-\alpha;\nu};$$

- se respinge H_0 în favoarea ipotezei H_1'' (în testul unilateral C), dacă:

$$t > t_{1-\alpha;\nu},$$

unde $\nu = n_1 + n_2 - 2$.

7.6.4. Testul Student pentru eşantioane independente, abateri standard inegale

Se menţin notaţiile de la testul Student precedent, pentru abateri standard egale, cu excepţia faptului că populaţia P_1 are o dispersie σ_1^2, iar populaţia P_2 o dispersie σ_2^2, dispersiile fiind diferite între ele.

Ipotezele testate privesc tot egalitatea mediilor celor două populaţii; prin urmare, se iau în considerare aceleaşi seturi de ipoteze (A), (B) sau (C).

Statistica testului devine, în forma simplificată de ipoteza nulă a egalităţii mediilor,

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

care are o repartiţie t cu numărul gradelor de libertate definit de relaţia

$$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\left(\frac{s_1^2}{n_1}\right)^2 \frac{1}{n_1+1} + \left(\frac{s_2^2}{n_2}\right)^2 \frac{1}{n_2+1}} - 2,$$

valoarea lui ν fiind, evident, rotunjită pentru a fi un număr întreg de grade de libertate.

Din punct de vedere formal, regula de decizie este aceeaşi ca la testul anterior.

7.6.5. Testul Z pentru eşantioane independente, dispersii cunoscute

În cazul în care dispersiile celor două populaţii normale, σ_1^2 şi σ_2^2, sunt cunoscute din studii anterioare, testarea egalităţii mediilor se efectuează cu ajutorul repartiţiei normale standard.

Dacă $\bar{x}_1$ şi $\bar{x}_2$ sunt mediile de sondaj obţinute din două eşantioane de volume n_1, respectiv n_2, atunci variabila

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}}$$

are o repartiţie normală standard, N(0;1), deci se pot testa diferenţe prezumtive între mediile populaţiilor.

Considerând, la fel ca la testele precedente, seturile de ipoteze (A), (B) sau (C), se observă că în condiţiile ipotezei nule de egalitate a mediilor, $\mu_1 = \mu_2$, statistica se simplifică prin dispariţia diferenţei dintre mediile populaţiilor:

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}}.$$

Regiunile critice ale testului sunt determinate, analog exemplului 7.2, de pe curba normală standard, astfel încât decizia în test presupune următoarele:

- se respinge H_0 în favoarea ipotezei H_1 (în testul bilateral A), dacă:

$$|Z| \geq z_{1-\alpha/2};$$

- se respinge H_0 în favoarea ipotezei H_1' (în testul unilateral B), dacă:

$$Z \leq -z_{1-\alpha} \text{ sau, echivalent, } Z \leq z_\alpha;$$

- se respinge H_0 în favoarea ipotezei H_1'' (în testul unilateral C), dacă:

$$Z \geq z_{1-\alpha}.$$

7.6.6. Testul pentru dou proporţii independente

Dacă variabila studiată este dihotomică, atunci este posibil ca, utilizând aproximarea normală a repartiţiei binomiale, să se compare două populaţii după proporţia cazurilor „succes".

Situaţia poate fi schematizată prin:

- două populaţii P_1 şi P_2, în care o variabilă aleatorie este repartizată binomial cu parametrii necunoscuţi p_1 şi, respectiv, p_2;

- două eşantioane, de volume n_1 şi n_2, extrase independent din cele două populaţii;
- în primul eşantion s-au întâlnit x_1 cazuri „succes", în al doilea eşantion s-au întâlnit x_2 cazuri „succes".

Reamintim că frecvenţa relativă x_1/n_1 este o estimaţie a parametrului p_1, iar x_2/n_2 este o estimaţie a parametrului p_2. Problema la care se caută o soluţie este să se determine o relaţie cât mai plauzibilă între parametrii p_1 şi p_2, relaţie care să fie susţinută de datele de sondaj disponibile.

Ipotezele testului sunt specifice unui test parametric:

$$\text{(A)} \begin{cases} H_0 : p_1 = p_2 \\ H_1 : p_1 \neq p_2 \end{cases}, \text{ pentru testul direcţional,}$$

$$\text{(B)} \begin{cases} H_0 : p_1 = p_2 \\ H_1' : p_1 < p_2 \end{cases} \text{ sau (C)} \begin{cases} H_0 : p_1 = p_2 \\ H_1'' : p_1 > p_2 \end{cases}, \text{ pentru test unilateral.}$$

Statistica testului este o statistică de tip Z, deci urmează o repartiţie normală standard. În condiţiile ipotezei nule, H_0, statistica are expresia:

$$Z = \frac{\left(\frac{x_1}{n_1} - \frac{1}{2n_1}\right) - \left(\frac{x_2}{n_2} + \frac{1}{2n_2}\right)}{\sqrt{p^*(1-p^*)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}},$$

unde sunt introduse, la numărător, corecţiile de continuitate, iar prin

$$p^* = \frac{x_1 + x_2}{n_1 + n_2}$$

se notează estimaţia valorii comune a parametrilor (în H_0, parametrii sunt egali, iar valoarea lor comună este estimată de p^*). Se observă, de altfel, că p^* se obţine ca fiind frecvenţa relativă a cazurilor „succes" în mulţimea formată prin reuniunea celor două eşantioane (reuniunea are sens în momentul în care presupunem că populaţiile sunt caracterizate de acelaşi parametru).

În cazul în care eşantioanele sunt suficient de mari, de regulă cu volume mai mari decât 30, variabila Z urmează o repartiţie normală standard şi, prin urmare, se poate decide în test după procedura generală a unui test Z:

- se respinge H_0 în favoarea ipotezei H_1 (în testul bilateral A), dacă:

$$|Z| \geq z_{1-\alpha/2};$$

- se respinge H_0 în favoarea ipotezei H_1' (în testul unilateral B), dacă:

$$Z \leq -z_{1-\alpha} \quad \text{sau, echivalent,} \quad Z \leq z_\alpha;$$

- se respinge H_0 în favoarea ipotezei H_1'' (în testul unilateral C), dacă

$$Z \geq z_{1-\alpha}.$$

Dacă dispunem de proporţiile cazurilor „succes" în cele două eşantioane, fie acestea r_1 şi r_2, atunci se pot regăsi frecvenţele absolute prin $x_1 = r_1 \times n_1$ şi $x_2 = r_2 \times n_2$, după care se aplică formulele anterioare, sau se poate utiliza direct formula

$$Z = \frac{\left(r_1 - \frac{1}{2n_1}\right) - \left(r_2 + \frac{1}{2n_2}\right)}{\sqrt{p^*(1-p^*)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}, \text{ cu } p^* = \frac{r_1 n_1 + r_2 n_2}{n_1 + n_2}.$$

Se observă, în această formă, că estimaţia proporţiei comune p^* este media celor două proporţii de sondaj ponderate cu volumele eşantioanelor.

Observaţia 1. Dacă volumele de sondaj sunt suficient de mari, atunci corecţiile de continuitate devin foarte mici şi nu mai influenţează rezultatele în primele zecimale. În acest caz, ele pot fi omise.

Observaţia 2. Dacă la numărătorul expresiilor care definesc statistica Z se scade cantitatea $(p_1 - p_2)$, atunci se poate utiliza variabila obţinută pentru a testa concordanţa diferenţei dintre parametri cu o valoare particulară a acestei diferenţe.

Exemplul 7.7. Elevilor unui liceu le-a fost aplicat un chestionar cu privire la şansa lor de reuşită în viaţă, fiind clasaţi fie idealişti, fie realişti. Răspunsurile celor din clasa a IX-a au fost comparate cu cele ale elevilor din clasa terminală. Cercetarea a cuprins două eşantioane independente; datele sunt prezentate în tabelul de contingenţă de mai jos (tabelul 7.5).

Se doreşte compararea proporţiilor de „idealişti" din cele două clase. Aplicăm un test bilateral cu

$\begin{cases} H_0: \text{proporţiile de idealişti sunt identice;} \\ H_1: \text{proporţiile de idealişti sunt diferite.} \end{cases}$

În spiritul dezbaterii generale a testului, se va considera „succes" înregistrarea unei clasări „idealist"; variabila este evident dihotomică etc.

Tabelul 7.5. *Datele pentru exemplul 7.7*

	clasa a IX-a	clasa a XII-a
idealişti	22 (62,9%)	16 (33,3%)
realişti	13 (37,1%)	32 (66,7%)
n	35	48

Utilizând proporţiile $r_1 = 0{,}629$ şi $r_2 = 0{,}333$, se obţine

$$p^* = \frac{0{,}629 \times 35 + 0{,}333 \times 48}{35 + 48} = 0{,}458,$$

care este estimaţia proporţiei comune a „idealiştilor". Pentru statistică se obţine valoarea

$$Z = \frac{\left(0{,}629 - \frac{1}{70}\right) - \left(0{,}333 + \frac{1}{96}\right)}{\sqrt{0{,}458 \times (1 - 0{,}458) \times \left(\frac{1}{35} + \frac{1}{48}\right)}} = \frac{0{,}271}{0{,}11} = 2{,}46,$$

care este mai mare decât $z_{0{,}975} = 1{,}96$; aşadar, se respinge ipoteza nulă a egalităţii proporţiilor de idealişti, la nivelul de semnificaţie $\alpha = 0{,}05$.

Dacă s-ar fi aplicat un test unilateral, de tip (C), concluzia ar fi fost că se respinge ipoteza nulă în favoarea ipotezei că proporţia de „idealişti" este mai mare în rândul elevilor de clasa a IX-a.

Testul statistic efectuat poate să pară inutil atâta timp cât, analizând datele de sondaj, se observă imediat că proporţiile sunt inversate în cele două populaţii. Această concluzie este nefondată, deoarece o modificare a datelor sub forma din tabelul 7.6, care susţine aceeaşi evidenţă la examinarea vizuală a tabelului, nu mai este semnificativă statistic.

Tabelul 7.6. *Datele modificate pentru exemplul 7.7*

	clasa a IX-a	clasa a XII-a
idealişti	21 (60,0%)	22 (45,8%)
realişti	14 (40,0%)	26 (54,2%)
n	35	48

Calculele efectuate pentru această configuraţie modificată a frecvenţelor, cu r_1=0,60 şi r_2=0,458, produc $p^* = 0{,}518$ şi $Z = 1{,}056$. La nivelul de semnificaţie $\alpha = 0{,}05$, nu se mai poate respinge ipoteza nulă; cu alte cuvinte, eşantioanele nu diferă atât de mult încât să susţină ipoteza că diferenţa se datorează deosebirilor dintre populaţii.

Cele două situaţii prezentate în exemplu vor să accentueze şi importanţa deosebită pe care trebuie să o acorde experimentatorul fiecărui rezultat experimental. Orice eroare introdusă în înregistrarea unei probe poate schimba semnificativ concluzia cercetării.

***

Următoarele teste se referă la eşantioane dependente. Sunt prezentate testele parametrice pentru variabile continue şi pentru variabile dihotomice. De remarcat, încă o dată, asemănarea care există între variabilele continue şi cele nominale dihotomice.

7.6.7. Testul Student pentru eşantioane cu date perechi

Reamintim că două eşantioane corelate (cu date perechi) se remarcă prin aceea că observaţiile pot fi grupate în perechi, câte o pereche pentru fiecare două elemente similare în cele două eşantioane.

Să luăm două populaţii caracterizate de o variabilă *normală* (sau aproape normală) *continuă* X. Dacă se compară cele două populaţii pe baza datelor de sondaj obţinute din două eşantioane corelate, atenţia cercetătorului este îndreptată asupra diferenţelor care apar între scorurile (măsurătorile) fiecărei perechi de observaţii. Pe baza acestor diferenţe este construit şi testul t pentru date perechi.

Datele se organizează în forma prezentată în tabelul 7.7, fiecare linie reprezentând o pereche de date „legate". De remarcat că două eşantioane perechi au acelaşi număr de elemente. Coloana diferenţelor marchează distribuţia empirică a unei noi variabile (variabila diferenţă) pentru care se pot calcula valorile tipice uzuale.

Tabelul 7.7. *Date perechi (corelate)*

X	X'	d (diferenţa)
X_1	X'_1	$d_1 = X_1 - X'_1$
X_2	X'_2	$d_2 = X_2 - X'_2$
...	...	...
X_n	X'_n	$d_n = X_n - X'_n$

Se calculează media şi abaterea standard s_d a diferenţelor prin

$$\bar{d} = \frac{1}{n}\sum d_i\,,\quad s_d^2 = \frac{1}{n-1}\sum (d_i - \bar{d})^2 = \frac{\sum d_i^2 - (\sum d_i)^2/n}{n-1}.$$

Notând, uzual, cu μ_1 şi μ_2, mediile celor două populaţii, ipotezele luate în considerare sunt:

(A) $\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases}$, pentru testul bilateral,

(B) $\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1' : \mu_1 < \mu_2 \end{cases}$ sau (C) $\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1'' : \mu_1 > \mu_2 \end{cases}$, pentru un test unilateral.

Statistica testului este:

$$t = \frac{\bar{d}}{s_d}\sqrt{n}\,.$$

Dacă ipoteza nulă este adevărată, atunci variabila d urmează o repartiţie normală $N(0;\sigma_d^2/n)$, iar variabila t urmează o repartiţie Student cu $\nu = n - 1$ grade de libertate.

Decizia în test se ia potrivit procedurii uzuale de la un test t:

- se respinge H_0 în favoarea ipotezei H_1 (în testul bilateral A), dacă:

$$|t| \geq t_{1-\alpha/2;\nu};$$

- se respinge H_0 în favoarea ipotezei H_1' (în testul unilateral B), dacă:

$$t < -t_{1-\alpha;\nu};$$

- se respinge H_0 în favoarea ipotezei H_1'' (în testul unilateral C), dacă:

$$t > t_{1-\alpha;\nu},$$

unde α este nivelul de semnificație fixat, iar $\nu = n - 1$.

Exemplul 7.8[50]. 15 subiecți au fost evaluați pe o scală de autoapreciere înainte și după intervenția unui factor experimental, urmărindu-se creșterea încrederii în sine. Rezultatele sunt prezentate în tabelul 7.8, împreună cu calculele auxiliare necesare aplicării testului Student.

Tabelul 7.8. *Datele de sondaj pentru exemplul 7.8*

Nr.crt.	Înainte	După	d	d^2
1	90	93	–3	9
2	91	90	1	1
3	93	89	4	16
4	89	88	1	1
5	85	88	–3	9
6	89	86	3	9
7	83	84	–1	1
8	88	83	5	25
9	84	83	1	1
10	82	80	2	4
11	83	77	6	36
12	81	76	5	25
13	72	74	–2	4
14	70	70	0	0
15	71	69	2	4
Total			**21**	**145**

Obținem, aplicând formulele anterioare:

$$\overline{d} = 1{,}4,\ s_d^2 = 8{,}26,\ s_d = 2{,}87,\ t = 1{,}887.$$

Pentru $\nu = 14$ grade de libertate și $\alpha = 0{,}05$ se citește, în tabelul repartiției Student, $t_{0{,}975;\,14} = 2{,}145$. Rezultă că nu se poate respinge ipoteza nulă în testul

50. *Apud* Norman R. Kurtz, *op.cit.*

bilateral și nu se mai verifică testul unilateral. Prin urmare, datele de sondaj nu susțin ipoteza că intervenția factorului experimental are influența așteptată.

> **Observație.** Deoarece suma diferențelor este pozitivă, este de așteptat că $\mu_2 > \mu_1$, adică experimentul la care au fost supuși subiecții are efectul scontat. Prin urmare, dacă ar trebui să se aplice un test unilateral, acesta ar fi doar testul notat cu (B).

7.6.8. Testul Z pentru două proporții dependente

Cercetarea unei variabile dihotomice necesită utilizarea unor tehnici particulare pentru compararea populațiilor implicate.

Să luăm cazul cel mai des întâlnit al situației „înainte/după". Datele pot fi aranjate sub forma din tabelul 7.9, unde, pentru simplitate, s-au notat cu „Da" și „Nu" valorile variabilei nominale dihotomice.

Tabelul 7.9. *Organizarea datelor pentru eșantioane dependente, variabila studiată fiind dihotomică*

Înainte/După	Da	Nu	Total
Da	a	b	a + b
Nu	c	d	c + d
Total	a + c	b + d	a + b + c + d = n

Interpretarea datelor este: *a* subiecți au răspuns Da înainte și după, *b* au răspuns Da înainte și Nu după, *c* au răspuns Nu înainte și Da după, iar *d* subiecți au răspuns Nu înainte și Nu după. Distribuțiile marginale au semnificații imediate: $a + b$ este numărul celor care au răspuns inițial Da, $c + d$ este numărul celor care au răspuns inițial Nu, $a+c$ este numărul celor care au răspuns Da după experiment, iar $b + d$ reprezintă numărul celor care au răspuns ulterior Nu.

Ceea ce se compară este structura răspunsurilor care s-a modificat între situația „înainte" și situația „după". Modificarea, dacă există, trebuie să fie evidențiată de numărul celor care își schimbă opțiunile: trec de la Da la Nu sau invers. Prin urmare modificarea de atitudine este ilustrată de căsuțele de pe diagonala stânga jos – dreapta sus, adică de numerele *c* și *b*.

A doua situație în care sunt implicate eșantioane dependente este cea în care dispunem efectiv de eșantioane distincte, cu elemente care corespund două câte două. Situația este ilustrată în figura 7.4.

Se observă astfel că elementele perechi sunt considerate echivalente, astfel încât *a*, similar tabelului 7.9, reprezintă numărul celor care și-au păstrat opțiunea Da, *b* este numărul celor care și-au schimbat atitudinea din Da în Nu etc.

Ipotezele se referă la proporțiile celor care răspund afirmativ în cele două populații (proporțiile cazurilor „succes", după terminologia de la variabilele binomiale):

$$\begin{cases} H_0: \text{proporția a + b este aceeași cu proporția a + c;} \\ H_1: \text{proporția a + b este diferită de proporția a + c.} \end{cases}$$

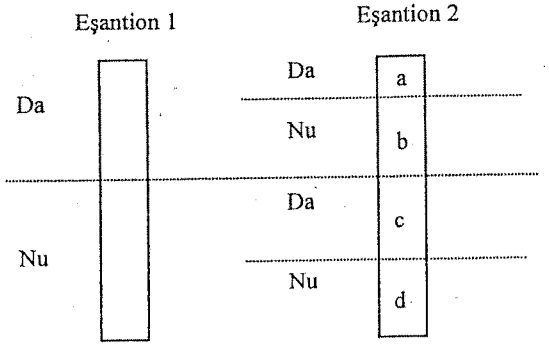


Figura 7.4. *Identificarea schimbărilor în eşantioane dependente distincte*

Este posibil să se ia în considerare şi unele ipoteze unilaterale, enunţul lor fiind evident.

Statistica testului se bazează pe valorile care arată schimbarea atitudinilor:

$$Z = \frac{|b-c|-1}{\sqrt{b+c}}.$$

În condiţiile ipotezei nule, statistica urmează o repartiţie normală standard. Prin urmare, pentru decizia în test, se poate urma procedura uzuală a unui test bazat pe repartiţia N(0;1), la un nivel de semnificaţie α fixat:

- se respinge ipoteza nulă dacă $|Z| \geq z_{1-\alpha/2}$;
- nu se respinge ipoteza nulă în caz contrar.

Exemplul 7.9. Într-un studiu privind fumatul, s-a selectat un eşantion de persoane care au fost supuse unui program de educare privind efectele nocive ale fumatului, program cu durata de 6 luni. Datele privind proporţiile de fumători, atât înainte de experiment, cât şi după terminarea programului, sunt cuprinse în tabelul 7.10:

Tabelul 7.10. *Datele de sondaj ale exemplului 7.9.*

Înainte/După	fumători	nefumători	Total
fumători	51	16	67
nefumători	7	36	43
Total	58	52	110

O analiză directă a tabelului de contingenţă arată că numărul celor care au renunţat la fumat este mai mare decât numărul celor care s-au „apucat" de fumat. Acest fapt se poate însă datora sondajului şi, pentru a decide dacă modificările sunt

atât de importante încât să poată fi atribuite programului educativ, este important să se aplice un test statistic. Ipotezele sunt:

$\begin{cases} H_0 : \text{proporţia de fumători este aceeaşi înainte şi după (program ineficient),} \\ H_1 : \text{proporţiile de fumători sunt diferite (program eficient).} \end{cases}$

Statistica testului este:

$$Z = \frac{|16-7|-1}{\sqrt{16+7}} = 1{,}668,$$

care nu permite respingerea ipotezei nule la nivelul $\alpha = 0{,}05$ deoarece $z_{0{,}975} = 1{,}96$.

Pentru a avea o imagine a sensibilităţii testului, dacă numărul celor care au abandonat fumatul ar fi de 18 (cu 2 mai mult) şi numărul celor care s-au apucat de fumat ar fi de 5 (cu doi mai puţin), atunci statistica ar fi de 2,50 şi s-ar putea respinge ipoteza nulă.

Observaţie. În aplicarea testului se recomandă ca fiecare diagonală a tabelului de contingenţă să aibă mai mult de 10 observaţii.

***

Testele prezentate în finalul acestui capitol se referă la mărimi discrete (nominale sau ordinale). Prezentarea se va face separat pentru eşantioane independente şi dependente. Fiind vorba de teste neparametrice, ele pot fi aplicate şi mărimilor continue, atunci când nu sunt asigurate condiţii de normalitate sau de efectiv, caz în care valorile vor fi grupate, în prealabil, în clase, potrivit procedurii uzuale.

7.6.9. Testul Mann-Whitney (U)

Atunci când este necesar să se decidă dacă două serii de date, referitoare la o anumită caracteristică, pot fi considerate dintr-o aceeaşi distribuţie, indiferent de forma acesteia, testul U poate constitui o metodă uşor de aplicat. Utilizarea testului Mann-Whitney este indicată, cu predilecţie, atunci când variabila după care se efectuează compararea populaţiilor este continuă, dar nu se poate decide asupra normalităţii ei. În această situaţie, nu se mai poate aplica testul Student şi s-a dezvoltat o procedură specifică, nu singura, de altfel.

Prin modul de aplicare, testul U poate fi utilizat şi pentru compararea după mărimi ordinale, cu relativ puţine valori determinate. Testul este denumit uneori test de verificare a omogenităţii.

Să considerăm cele două serii de date independente: $\{x_{11}, x_{12}, x_{13}, \ldots, x_{1n}\}$ şi $\{x_{21}, x_{22}, x_{23}, \ldots, x_{2m}\}$. Evident, ele sunt gândite ca provenind din două eşantioane independente de volum n, respectiv m.

Ipoteza nulă este aceea a identităţii repartiţiilor. Dacă variabila este continuă, atunci identitatea se reflectă în egalitatea mediilor, medianelor etc. În acest caz, se poate considera că testul este unul de comparare a valorilor medii. Practica arată,

de altfel, că testul U este mai sensibil la diferenţe ale mediilor decât la diferenţe ale dispersiilor.

Ipoteza alternativă este nedirecţională şi afirmă că repartiţiile sunt distincte în cele două populaţii.

Atribuirea de ranguri. Numim **rang** al unei valori x, cu privire la un ansamblu de valori A care conţine x, locul pe care-l ocupă x atunci când A este ordonat (crescător sau descrescător). Dacă elementele lui A sunt toate distincte, atunci operaţiunea este evidentă şi produce un rezultat unic. În cazul în care A are mai multe elemente egale cu x, atunci lui x i se atribuie rangul obţinut ca medie aritmetică a rangurilor care li se cuvin elementelor egale cu x. Această procedură este cunoscută ca desemnând atribuirea de ranguri cu tratarea cazurilor de *ex aequo* (de egalitate).

Exemplul 7.10. Pentru şirul 2, 3, 1, 6, 3, 8, 6, 3, ordonarea produce 1, 2, 3, 3, 3, 6, 6, 8, deci 8 valori dintre care 3 apare de trei ori, iar 6 de două ori. Atribuirea iniţială a rangurilor, ca şi cum toate valorile ar fi distincte:

valoare:	1	2	3	3	3	6	6	8
rang iniţial:	1	2	3	4	5	6	7	8.

Tratarea cazurilor de *ex aequo* produce rangurile finale:

valoare:	1	2	3	3	3	6	6	8
rang final:	1	2	4	4	4	6,5	6,5	8,

unde, de exemplu, cele două valori 6 aveau rangurile iniţiale 6 şi 7, media fiind 6,5, care devine rangul final al fiecărei valori 6.

> **Observaţie.** Este cunoscută şi procedura prin care rangul atribuit este primul rang ce revine unei valori dintr-un subşir de valori identice.

Statistica testului se obţine parcurgând următoarele etape:

a) cele n + m valori se consideră ca fiind împreună şi se ordonează crescător;
b) se atribuie ranguri valorilor ordonate ale şirurilor reunite (după procedura explicată anterior, cu tratarea cazurilor de *ex aequo*);
c) se identifică rangurile valorilor aparţinând fiecărui eşantion; notăm cu T_1 suma rangurilor pentru primul eşantion şi cu T_2 suma rangurilor pentru al doilea eşantion;
d) se calculează cantităţile:

$$W_1 = n \cdot m + \frac{n(n+1)}{2} - T_1,$$

$$W_2 = n \cdot m + \frac{m(m+1)}{2} - T_2;$$

e) statistica testului este valoarea minimă dintre W_1 şi W_2:

$$U = \min(W_1, W_2).$$

Decizia în test presupune următoarele:

- se respinge ipoteza nulă a identităţii repartiţiilor, la nivelul de semnificaţie α fixat, dacă valoarea U calculată este mai mică sau egală cu valoarea critică tabelată $W_{1-\alpha;n;m}$; valorile critice sunt date în tabelul 7.11;
- nu se respinge ipoteza nulă în caz contrar.

Tabelul 7.11 conţine valorile critice ale testului Mann-Whitney pentru valori ale efectivelor eşantioanelor până la 20 şi valorile uzuale ale lui α. Pentru valori mai mari decât 20, se poate utiliza aproximarea normală:

$$W_{1-\alpha;n;m} = \frac{n \cdot m}{2} - z_{1-\alpha/2}\sqrt{\frac{n \cdot m \cdot (n+m+1)}{12}},$$

ceea ce presupune utilizarea, drept statistică a testului, a expresiei

$$Z = \frac{U - \frac{n \cdot m}{2}}{\sqrt{\frac{n \cdot m \cdot (n+m+1)}{12}}}.$$

Tabelul 7.11. *Valorile critice $W_{1-\alpha;n;m}$ pentru testul Mann-Whitney*

m/n	α	5	6	8	10	12	14	16	18	20
3	**0,05**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
	0,01	–	–	–	0	1	1	2	2	3
4	**0,05**	**1**	**2**	**4**	**5**	**7**	**9**	**11**	**12**	**14**
	0,01	–	0	1	2	3	4	5	6	8
5	**0,05**	**2**	**3**	**6**	**8**	**11**	**13**	**15**	**18**	**20**
	0,01	0	1	2	4	6	7	9	11	13
6	**0,05**	**3**	**5**	**8**	**11**	**14**	**17**	**21**	**24**	**27**
	0,01	1	2	4	6	9	11	13	16	18
8	**0,05**	**6**	**8**	**13**	**17**	**22**	**26**	**31**	**36**	**41**
	0,01	2	4	7	11	15	18	22	26	30
10	**0,05**	**8**	**11**	**17**	**23**	**29**	**36**	**42**	**48**	**55**
	0,01	4	6	11	16	21	26	31	37	42
12	**0,05**	**11**	**14**	**22**	**29**	**37**	**45**	**53**	**61**	**69**
	0,01	6	9	15	21	27	34	41	47	54
14	**0,05**	**13**	**17**	**26**	**36**	**45**	**55**	**64**	**74**	**83**
	0,01	7	11	18	26	34	42	50	58	67
16	**0,05**	**15**	**21**	**31**	**42**	**53**	**64**	**75**	**86**	**98**
	0,01	9	13	22	31	41	50	60	70	79
18	**0,05**	**18**	**24**	**36**	**48**	**61**	**74**	**86**	**99**	**112**
	0,01	11	16	26	37	47	58	70	81	92
20	**0,05**	**20**	**27**	**41**	**55**	**69**	**83**	**98**	**112**	**127**
	0,01	13	18	30	42	54	67	79	92	105

Utilizarea acestei aproximări se face după procedura obişnuită a testelor bazate pe repartiţia normală standard.

Testul poate fi aplicat şi pentru eşantioane de volume relativ mici.

Exemplul 7.11. Un număr de elevi din două licee, selectaţi în două eşantioane independente, au fost supuşi unui test privind cunoştinţele de cultură generală. Scorul obţinut este pe o scală 0 – 40. Rezultatele sunt:

eşantion 1	18, 25, 37, 21, 16, 35, 28, 10
eşantion 2	15, 35, 28, 21, 21, 38, 29, 30, 16, 24

Pentru a calcula statistica testului U, parcurgem etapele descrise în prezentarea testului:

a) ordonăm toate valorile într-un singur şir; valorile primului eşantion sunt subliniate:
   10, 15, 16, 16, 18, 21, 21, 21, 24, 25, 28, 28, 29, 30, 35, 35, 37, 38;
b) se atribuie ranguri, cu tratarea cazurilor de *ex aequo*:
   1; 2; 3,5; 3,5; 5; 7; 7; 7; 9; 10; 11,5; 11,5; 13; 14; 15,5; 15,5; 17; 18;
c) se calculează $T_1 = 1+3,5+5+7+10+11,5+15,5+17 = 70,5$ pentru suma rangurilor valorilor din primul eşantion şi, în mod analog, $T_2 = 100,5$ pentru rangurile valorilor din eşantionul al doilea;
d) se calculează, $n = 8$ şi $m = 10$, cantităţile $W_1 = 45,5$ şi $W_2 = 34,5$;
e) statistica testului este $U = \min(45,5; 34,5) = 34,5$.

Cum valoarea critică, în tabelul 7.11, este $W_{0,95;\,8;\,10} = 17$, rezultă că $U > W_{0,95;\,8;\,10}$ şi deci nu se poate respinge ipoteza nulă a identităţii celor două repartiţii. Deci, deşi eşantionul format din elevii liceului al doilea are acumulate mai multe puncte în medie, nu se susţine statistic ipoteza că elevii liceului al doilea au o cultură generală mai bună.

Observaţii. 1) Testul Mann-Whitney este cea mai bună alternativă la testul Student pentru compararea datelor continue.
2) Aplicarea testului nu este recomandată pentru cazuri multiple de *ex aequo*. Respectiva situaţie poate să apară frecvent dacă se utilizează o scală ordinală cu puţine categorii şi dacă eşantioanele au volume mari. În aceste situaţii se aplică testul Kolmogorov-Smirnov pentru două eşantioane, prezentat în continuare.

7.6.10. Testul Kolmogorov-Smirnov (K-S) pentru dou eşantioane

Pentru date ordinale cu multe situaţii de egalitate, ca şi pentru date continue grupate în relativ puţine clase, compararea repartiţiilor a două populaţii se poate efectua printr-un test Kolmogorov-Smirnov dezvoltat pentru cazul a două eşantioane.

Se consideră două eşantioane independente, de volume n_1 şi, respectiv, n_2. Pe baza datelor de sondaj corespunzătoare unei variabile ordinale, se doreşte

compararea populaţiilor pentru care eşantioanele sunt reprezentative. Testul poate fi aplicat dacă atât n_1, cât şi n_2 sunt mai mari decât 40.

Modul de calcul al statisticii testului este similar testului K-S pentru un eşantion (testul de concordanţă). Se pleacă de la tabelul statistic al frecvenţelor relative cumulate (vezi tabelul 7.12) şi se determină diferenţa maximă dintre frecvenţele corespunzătoare celor două eşantioane.

Tabelul 7.12. *Explicaţii la testul Kolmogorov-Smirnov (două eşantioane)*

Categoria	Eşantion 1		Eşantion 2		Diferenţa
	frecvenţa absolută	frecvenţa relativă cumulată	frecvenţa absolută	frecvenţa relativă cumulată	
c_1	f_1	$F^*(c_1)$	g_1	$G^*(c_1)$	$\lvert F^*(c_1) - G^*(c_1)\rvert$
c_2	f_2	$F^*(c_2)$	g_2	$G^*(c_2)$	$\lvert F^*(c_2) - G^*(c_2)\rvert$
...	...	...	...	...	...
c_k	f_k	$F^*(c_k)$	g_k	$G^*(c_k)$	$\lvert F^*(c_k) - G^*(c_k)\rvert$

Statistica testului Kolmogorov-Smirnov pentru două eşantioane este

$$D = \max_i \left| F^*(c_i) - G^*(c_i) \right|.$$

Testul poate fi utilizat atât în varianta bilaterală, cât şi în varianta unilaterală. Regula de decizie diferă în cele două ipostaze. Ipoteza nulă, chiar dacă pare a se referi la eşantioane, trebuie înţeleasă ca aplicându-se pentru repartiţia variabilei în populaţii.

Testul bilateral

Ipotezele sunt:

- H_0 : diferenţa maximă dintre distribuţiile de frecvenţe cumulate este egală cu zero (deci distribuţiile sunt identice);
- H_1 : diferenţa maximă dintre distribuţiile de frecvenţe cumulate este diferită de zero (deci distribuţiile sunt diferite).

Decizia în test se bazează pe compararea statisticii calculate cu valoarea critică ce corespunde nivelului de semnificaţie şi volumelor celor două eşantioane. Tabelul 7.13 arată valorile critice ale testului K-S pentru două eşantioane. Sunt luate în considerare principalele valori ale nivelului de semnificaţie.

Se respinge ipoteza nulă dacă valoarea calculată este mai mare sau egală cu valoarea critică D* corespunzătoare.

Tabelul 7.13. *Valorile critice ale testului K-S pentru două eşantioane*

α	valoarea critică (D*)
0,10	$1{,}22\sqrt{\frac{n_1+n_2}{n_1 n_2}}$
0,05	$1{,}36\sqrt{\frac{n_1+n_2}{n_1 n_2}}$
0,025	$1{,}48\sqrt{\frac{n_1+n_2}{n_1 n_2}}$
0,01	$1{,}63\sqrt{\frac{n_1+n_2}{n_1 n_2}}$
0,005	$1{,}73\sqrt{\frac{n_1+n_2}{n_1 n_2}}$
0,001	$1{,}95\sqrt{\frac{n_1+n_2}{n_1 n_2}}$

Testul unilateral

Reamintim că pentru datele ordinale are sens să se vorbească despre forma şi localizarea distribuţiei. În testul direcţional se va decide asupra poziţiilor relative ale celor două distribuţii, ipotezele fiind:

- H_0 : diferenţa maximă dintre distribuţiile de frecvenţe cumulate este egală cu zero (distribuţiile sunt identice);
- H'_1: diferenţa maximă este diferită de zero şi arată că o distribuţie (aceea care produce diferenţa maximă pozitivă) este situată la dreapta celeilalte.

Statistica testului este, în condiţiile ipotezei nule, o variabilă repartizată χ^2 cu $\nu = 2$ grade de libertate:

$$\chi^2 = 4D^2 \times \frac{n_1 n_2}{n_1 + n_2}.$$

Regula de decizie este atunci în funcţie de următoarele:

- se respinge ipoteza nulă în favoarea ipotezei alternative H'_1 dacă valoarea calculată depăşeşte quantila corespunzătoare a repartiţiei χ^2:

$$\chi^2 \geq \chi^2_{1-\alpha;2};$$

- nu se respinge ipoteza nulă, în caz contrar.

Se poate observa că, datorită numărului de grade de libertate independent de volumele eşantioanelor disponibile, decizia în test este simplă. De exemplu, pentru $\alpha = 0{,}05$ se poate reţine valoarea quantilei $\chi^2_{0{,}95;2} = 4{,}60$ şi, prin urmare, decizia la nivelul de siguranţă statistică de 95% se obţine comparând valoarea calculată cu 4,60, indiferent de volumele eşantioanelor.

Exemplul 7.12. Într-un studiu comparativ care implică studierea repartiţiilor unei caracteristici ordinale în două zone geografice, s-au obţinut datele de sondaj din tabelul 7.14. Tabelul este completat cu frecvenţele relative cumulate (vezi tabelul 7.12) şi diferenţele rezultate.

Datorită numărului mare de egalităţi, nu se poate aplica testul U. Astfel, există 25 + 31 cazuri de opinii „foarte favorabil", ceea ce în ordonarea impusă de testul Mann-Whitney produce 56 de ranguri egale cu 28,5 etc.

Tabelul 7.14. *Datele de sondaj ale exemplului 7.12.*

Categoria	Zona 1		Zona 2		Diferenţa
	frecvenţa absolută	frecvenţa relativă cumulată	frecvenţa absolută	frecvenţa relativă cumulată	
foarte favorabil	25	0,124	31	0,130	0,006
favorabil	42	0,333	50	0,340	0,007
indiferent	64	0,652	84	0,693	0,041
nefavorabil	52	0,910	48	0,895	0,015
foarte nefavorabil	18	1,000	25	1,000	0
Total	**201**		**238**		

Se obţine $D = 0{,}041$, diferenţa maximă absolută a celor două distribuţii de sondaj cumulate. Aplicând, pentru $\alpha = 0{,}05$, formula pentru valoarea critică D*, din tabelul 7.13 rezultă:

$$D^* = 1{,}36\sqrt{\frac{201+238}{201\times 238}} = 0{,}130.$$

Cum valoarea calculată este mai mică decât valoarea critică, nu putem respinge ipoteza nulă a identităţii repartiţiilor. De remarcat că analiza directă a frecvenţelor relative cumulate poate să sugereze o mai mare frecvenţă a opiniilor favorabile sau indiferente în zona a doua.

7.6.11. Testul Wilcoxon

Acest test este caracterizat ca fiind un test de comparare pentru mărimi ordinale dependente, deşi se aplică pentru date de interval. Respectiva apreciere se datorează faptului că statistica testului se calculează utilizând ranguri, deci o procedură specifică datelor ordinale.

Testul Wilcoxon este utilizat pentru a determina dacă două eşantioane dependente reprezintă populaţii similare. Este un test neparametric, deci poate fi utilizat atunci când aplicarea unui test Student pentru date perechi nu se poate efectua datorită neîndeplinirii condiţiilor de normalitate.

Statistica testului se bazează pe studiul diferenţelor care apar între observaţiile perechi, considerându-se atât semnul diferenţelor, cât şi mărimea lor absolută.

Ipotezele testate sunt, în general, de tip bilateral, dar se utilizează şi ipoteze alternative direcţionale.

Testul bilateral:

- H_0 : distribuţiile variabilei în cele două populaţii sunt identice;
- H_1 : distribuţiile variabilei în cele două populaţii sunt diferite.

După semnul diferenţelor observate, se poate formula, pentru un test unilateral, una dintre ipotezele alternative:

- H_1' : distribuţiile variabilei în prima populaţie se află la stânga distribuţiei din a doua populaţie (diferenţe negative),

sau

- H_1'' : distribuţia variabilei în prima populaţie este situată la dreapta distribuţiei din a doua populaţie (diferenţe pozitive)

Statistica testului se obţine prin parcurgerea următoarelor etape:

a) se calculează diferenţele observate:

eşantion 1 (pre-test)	eşantion 2 (post-test)	diferenţa
x_1	y_1	x_1-y_1
x_2	y_2	x_2-y_2
...	...	...
x_n	y_n	x_n-y_n

b) se elimină diferenţele nenule (cele care nu arată schimbări între observaţiile corespunzătoare); se obţine şirul diferenţelor nenule

$$d_1, d_2, \ldots, d_k,$$

unde, evident, $k \leq n$;

c) se ordonează şirul diferenţelor nenule absolute

$$|d_{(1)}| \leq |d_{(2)}| \leq \ldots \leq |d_{(k)}|,$$

unde notaţiile sunt cele consacrate pentru statistici de ordine;

d) se atribuie ranguri şirului obţinut la c), cu tratarea cazurilor de *ex aequo* (vezi procedura de la testul Mann-Whitney);

e) se calculează suma rangurilor care corespund diferenţelor iniţiale negative (notată cu T_-) şi suma rangurilor care corespund diferenţelor iniţiale pozitive (notată cu T_+);

f) statistica testului, notată cu T, este valoarea minimă a celor două sume calculate la e): $T = \min(T_-, T_+)$.

Decizia în test priveşte următoarele situaţii:

- se respinge ipoteza H_0 în favoarea ipotezei H_1 (testul bilateral) dacă valoarea calculată T este mai mică sau egală *cu valoarea critică tabelată pentru testul bilateral*;
- se respinge ipoteza H_0 în favoarea ipotezei alternative H_1' (sau H_1'') dacă valoarea calculată T este mai mică sau egală cu *valoarea critică tabelată pentru testul unilateral*.

Valorile critice pentru testul Wilcoxon sunt tabelate (vezi tabelul 7.15) pentru valori mici ale lui *n* (sub 25). În cazul valorilor mari ale volumului eşantioanelor, se poate recurge la aproximarea normală:

$$Z = \frac{T - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}.$$

Decizia în test se va lua atunci după procedura unui test bazat pe repartiţia normală standard prin compararea statisticii calculate cu quantila $z_{1-\alpha/2}$ (testul bilateral) sau cu $\pm z_{1-\alpha}$ (în testele unilaterale).

Tabelul 7.15. *Valorile critice T* pentru testul Wilcoxon*

	Nivelul de semnificaţie pentru testul unilateral		
	0,025	0,01	0,005
	Nivelul de semnificaţie pentru testul bilateral		
n	0,05	0,02	0,01
6	0	–	–
7	2	0	–
8	4	2	0
9	6	3	2
10	8	5	3
11	11	7	5
12	14	10	7
13	17	13	10
14	21	16	13
15	25	20	16
16	30	24	20
17	35	28	23
18	40	33	28
19	46	38	32
20	52	43	38
21	59	49	43
22	66	56	49
23	73	62	55
24	81	69	61
25	89	77	68

Exemplul 7.13. Într-o cercetare privind efectele unei noi metode de învățare a unei limbi străine, s-au selectat două eşantioane corelate. Primul eşantion a învățat după metodele obişnuite (lotul martor), iar celălalt după metoda nouă (lotul experimental). După o perioadă de timp, a fost înregistrat numărul de cuvinte noi memorate de cele două loturi. Se caută răspuns la întrebarea dacă metoda propusă este mai eficientă (conduce la memorarea unui număr crescut de cuvinte în aceeaşi perioadă de timp).

Datele de sondaj sunt conținute în tabelul 7.16, fiind însoțite şi de prelucrările necesare aplicării testului Wilcoxon. De remarcat că variabila implicată (număr de cuvinte) nu este o variabilă continuă pură (este greu de definit ce înseamnă memorarea unei fracțiuni de cuvânt) şi se apropie mai mult de o variabilă ordinală. Aplicarea unui test Student pentru date perechi este, din acest motiv, oarecum forțată. Analiza directă a datelor observate poate să ducă doar la opinia că metoda nouă este mai bună, deoarece există atât diferențe negative favorabile, cât şi diferențe pozitive, care sunt defavorabile.

Tabelul 7.16. *Datele de sondaj pentru exemplul 7.13*

nr. crt.	lot martor	lot experimental	diferența	diferența absolută	rangul diferenței	marcaj (+/–)
1	76	82	–6	6	11,5	–
2	79	80	–1	1	2	–
3	65	66	–1	1	2	–
4	82	78	4	4	7,5	+
5	75	83	–8	8	15	–
6	68	72	–4	4	7,5	–
7	57	64	–7	7	13,5	–
8	78	78	0	0	–	
9	83	90	7	7	13,5	+
10	63	68	–5	5	9,5	–
11	77	71	6	6	11,5	+
12	61	70	–9	9	16,5	–
13	66	71	–5	5	9,5	–
14	73	75	–2	2	4,5	–
15	82	82	0	0	–	
16	91	90	1	1	2	+
17	86	89	–3	3	6	–
18	75	90	–15	15	18	–
19	80	71	9	9	16,5	+
20	77	79	–2	2	4,5	–

Calculând statisticile T_+ şi T_-, se obțin, pe baza informațiilor din ultimele două coloane ale tabelului, valorile $T_+ = 51$ şi $T_- = 120$. Rezultă atunci că statistica Wilcoxon este $T = 51$, adică valoarea minimă dintre cele două calculate anterior.

Tabelul valorilor critice arată pentru testul bilateral, la volumul eşantionului $n = 20$, valoarea $T^* = 52$ corespunzătoare nivelului de semnificație $\alpha = 0{,}05$. Cum valoarea calculată este mai mică decât valoarea critică, rezultă că se poate respinge

ipoteza nulă a identității repartițiilor. În acest caz, se poate continua cu testul unilateral şi se poate respinge, la nivelul $\alpha = 0{,}025$, ipoteza nulă în favoarea ipotezei că distribuția în cea de-a doua populație este mai la dreapta; altfel spus noua metodă duce la memorarea mai multor cuvinte în acelaşi interval de timp. De remarcat că examinarea directă a rezultatelor nu susține o concluzie foarte fermă atâta timp cât diferențele în favoarea primei metode sunt printre cele mai mari în valoare absolută.

7.6.12. Testul McNemar

Situația caracteristică pentru aplicarea testului McNemar este:

- variabilă dihotomică,
- eşantioane dependente.

Se observă că acestea sunt şi condițiile de aplicare a unui test de comparare a proporțiilor, testul McNemar fiind o alternativă la respectivul test.

De obicei, testul este aplicat în situații pre-test/post-test, un tabel de date uzual prezentându-se sub forma:

	după	
înainte	**nu**	**da**
da	*a*	*b*
nu	*c*	*d*

În acest tabel, valorile *a* şi *d* măsoară schimbarea care a avut loc în urma experimentului şi se caută semnificația trecerilor de la „da" la „nu" (adică *a*) față de trecerile de la „nu" la „da" (adică *d*).

Ipotezele testului sunt:

- H_0 : schimbarea nu este semnificativă;
- H_1 : schimbarea este semnificativă.

Este evident că aceste ipoteze pot fi formulate şi în termeni de repartiții ale populațiilor respective.

Deşi ipotezele sunt nedirecționale, testul McNemar este în esență un test direcțional care priveşte modificarea de la *a* la *d*.

Statistica testului este similară celei de la testul de comparare a proporțiilor, dar se bazează pe repartiția χ^2:

$$\chi^2 = \frac{(|a-d|-1)^2}{a+d}.$$

În condițiile ipotezei nule, statistica are o repartiție χ^2 cu un grad de libertate, astfel încât decizia este următoarea:

- se respinge ipoteza nulă în favoarea ipotezei alternative, la nivelul de semnificație α, dacă valoarea calculată este mai mare decât quantila corespunzătoare a repartiției χ^2:

$$\chi^2 \geq \chi_{1;\,1-\alpha}.$$

Testul McNemar este neparametric, dar trebuie ca volumul eşantioanelor să fie suficient de mare încât să asigure observarea unui număr de schimbări mai mare decât 10 ($a + d > 10$).

7.6.13. Testul χ^2 de comparare (asociere)

Dezvoltat pentru date nominale, testul χ^2 este unul dintre cele mai utilizate teste în ştiinţele sociale, atât datorită specificului datelor experimentale din aceste domenii, cât şi datorită uşurinţei cu care se efectuează calculele şi se ia decizia în acest test.

Situaţiile de aplicare a testului sunt şi ele multiple. Dacă sunt implicate două populaţii (prin eşantioane reprezentative), atunci testul poate fi considerat *test de comparare*: se compară repartiţia variabilei studiate în cele două populaţii.

Dacă sunt implicate mai multe populaţii diferenţiate după categoriile unei variabile, cum ar fi *locul de reşedinţă* cu valorile urban, suburban, rural, atunci testul χ^2 poate fi gândit ca un *test de asociere*. Dacă notăm variabila studiată cu X, iar variabila care diferenţiază subpopulaţiile cu Y, atunci decizia în test poate privi modul de asociere a variabilelor X şi Y, adică modul în care o modificare a valorilor lui Y se reflectă într-o modificare a valorilor lui X.

Diferenţierea între cele două moduri de interpretare a deciziei în test va fi făcută acolo unde este cazul. Limbajul adoptat pentru prezentare este însă cel al asocierii dintre variabile (a se vedea şi dezbaterea din capitolul dedicat asocierii variabilelor).

Aplicarea testului cere ca datele să fie sub formă de frecvenţe absolute, adică variabilele să fie, de preferinţă, variabile categoriale sau ordinale. În cazul când se aplică unor variabile continue, ele vor fi discretizate prin gruparea pe intervale în funcţie de unele praguri acceptate. De exemplu: valorile normale ale variabilei sunt cele mai mari decât 12, restul valorilor fiind anormale; în acest mod, dintr-o variabilă continuă s-a obţinut una cu două categorii: normal/anormal.

Presupunem că poate prezenta interes studierea asocierii dintre variabilele X şi Y. Presupunem că variabila X se măsoară pe o scală cu m categorii, fie ele $x_1, x_2, \ldots, x_m$, iar variabila Y se măsoară pe o scală cu n categorii, fie ele $y_1, y_2, \ldots, y_n$. Deoarece *variabilele* au fost măsurate simultan pentru toate observaţiile (fiecare observaţie are o valoare pentru X şi una pentru Y), dispunem de un tabel de contingenţă clasic. Notăm cu N numărul total de observaţii.

Frecvenţele observate

X/Y	y_1	y_2	...	y_n	Total (pe linie)
x_1	O_{11}	O_{12}	...	O_{1n}	O_{1*}
x_2	O_{21}	O_{22}	...	O_{2n}	O_{2*}
...	...	...	...	...	...
x_m	O_{m1}	O_{m2}	...	O_{mn}	O_{m*}
Total (pe coloană)	O_{*1}	O_{*2}	...	O_{*n}	N

Prin O_{ij} s-a notat numărul **observat** de cazuri care se încadrează în nivelul i după X şi în nivelul j după Y. Prin totalul O_{i*} s-a notat totalul de pe linia i, iar prin O_{*j} s-a notat totalul de pe coloana j.

Întrebarea la care se cere răspuns este dacă variabilele X şi Y sunt asociate, adică dacă apariţia unui nivel al lui X favorizează apariţia unor anumite niveluri ale lui Y (sau reciproc). Alternativa este aceea în care frecvenţele observate ar fi rezultatul întâmplării. Se obişnuieşte astfel să se considere un tabel, similar ca organizare celui observat, cu frecvenţele aşteptate. Aici, sensul termenului *aşteptat* ţine de *ceea ce ar trebui să observăm în cazul în care între X şi Y nu există asociere*; aşadar, configuraţiile de frecvenţe trebuie să fie identice pe linii şi pe coloane cu configuraţiile marginale, adică distribuţia de la nivelul general se regăseşte pe fiecare linie/coloană (cu alte cuvinte, comportarea valorilor nu depinde de linie/coloană).

Frecvenţele aşteptate

X/Y	y_1	y_2	...	y_n	Total (pe linie)
x_1	E_{11}	E_{12}	...	E_{1n}	O_{1*}
x_2	E_{21}	E_{22}	...	E_{2n}	O_{2*}
...	...	...	...	...	...
x_m	E_{m1}	E_{m2}	...	E_{mn}	O_{m*}
Total (pe coloană)	O_{*1}	O_{*2}	...	O_{*n}	N

Se observă că totalurile se păstrează, deoarece este vorba de aceleaşi efective distribuite altfel. Frecvenţele aşteptate sunt calculate, potrivit logicii de repartiţie similară pe linii şi coloane, după formula:

$$E_{ij} = \frac{\text{Total pe linie} \times \text{Total pe coloană}}{\text{Total general}} = \frac{O_{i*} \times O_{*j}}{N}.$$

Se va decide că X şi Y sunt asociate, dacă între cele două tabele există o mare diferenţă de distribuţie a frecvenţelor. Se măsoară diferenţa dintre tabele prin cantitatea, denumită χ^2:

$$\chi^2 = \sum_{i,j} \frac{(O_{ij} - E_{ij})^2}{E_{ij}},$$

suma fiind extinsă asupra tuturor căsuţelor din tabel (toate valorile i şi j).

Ipotezele testului de asociere χ^2 sunt:

$$\begin{cases} H_0: \text{variabilele } X \text{ şi } Y \text{ nu sunt asociate;} \\ H_1: \text{variabilele } X \text{ şi } Y \text{ sunt asociate.} \end{cases}$$

Se demonstrează că, în condiţiile ipotezei H_0, variabila χ^2, calculată mai sus, este repartizată χ^2 cu un număr de grade de libertate, v, egal cu $(m-1)\times(n-1)$, adică $\chi^2 \sim \chi^2_{(n-1)(m-1)}$, ceea ce permite efectuarea unui test statistic.

Decizia în test se poate lua:

- pe baza comparării valorii calculate χ^2 cu valoarea critică tabelată corespunzătoare pragului de semnificație α fixat; se respinge ipoteza nulă dacă valoarea calculată este mai mare decât valoarea critică tabelată;
- pe baza interpretării probabilității critice a testului; ipoteza nulă se respinge dacă probabilitatea critică ce a fost calculată este mai mică decât pragul de semnificație α fixat.

În cazul în care nu se respinge ipoteza nulă, nu înseamnă că între cele două variabile nu există asociere. Trebuie interpretat că *asocierea nu a putut fi pusă în evidență de datele existente* sau că *asocierea nu este statistic semnificativă.*

Prin respingerea ipotezei nule se afirmă doar existența asocierii dintre variabile, nu și măsurarea gradului de asociere. Pentru aceasta se utilizează alte tehnici statistice, de exemplu coeficienții de corelație.

Limite de aplicare. Pentru a aplica testul trebuie ca nu mai mult de 20% dintre casete să aibă mai puțin de 5 elemente.

***

Exemplele care urmează au mai ales rolul de a evidenția cazurile principale de aplicare a testului χ^2. Fiecare exemplu va fi, prin urmare, însoțit de un comentariu asupra specificității situației analizate.

Exemplul 7.14 (modelul 2 × 2). Compararea a două populații după o variabilă dihotomică s-a efectuat atât în cazul eșantioanelor dependente, cât și în cazul eșantioanelor independente, prin luarea în considerare a proporțiilor. Problema poate fi privită și din alt punct de vedere, dacă se consideră că populațiile implicate sunt subpopulații ale unei populații mai vaste, subdiviziuni definite de o nouă variabilă dihotomică (cum ar fi bărbat/femeie, adult/copil, urban/rural etc.). În acest caz se poate vorbi de asocierea dintre variabila studiată și variabila de definire a subpopulațiilor.

Un studiu privind frecvența persoanelor care fumează a condus la următorul tabel de date:

Tabelul 7.17. *Datele observate în exemplul 7.14*

	bărbați	femei	total
fumători	25	26	51
nefumători	18	38	56
total	43	64	107

Analiza directă a datelor poate conduce la concluzii de genul: există mai multe femei care fumează decât bărbați care fumează (deoarece 28 > 23) sau: proporția bărbaților fumători este mai mare decât proporția femeilor fumătoare. Concluziile de acest gen au marele defect că se referă strict la eșantion și nu constituie concluzii aplicabile la populațiile implicate. *Asemenea concluzii pot să constituie doar suportul pentru emiterea unor ipoteze, verificarea acestora trebuind să se facă pe baza datelor de sondaj din alte eșantioane. Verificarea pe baza datelor de sondaj care au constituit suportul de formare a ipotezei nu este, din motive evidente, o metodă științifică acceptabilă.* Din acest motiv, vom presupune că ipotezele testate sunt apriorice culegerii de date sau, cu alte cuvinte, ipotezele considerate nu au influențat, voit sau nu, obținerea datelor. Necesitatea aparatului statistic este deci esențială și în asemenea structuri foarte simple de date.

Tabelul 7.17 poate fi gândit ca opunând o populație de sex masculin unei populații de sex feminin. În acest caz, se va decide dacă, din punctul de vedere al fumatului, cele două populații sunt identice sau nu. Datele au fost obținute în acest caz prin considerarea a două eșantioane independente (unul masculin și unul feminin) și prin determinarea proporțiilor de fumători/nefumători în cele două eșantioane.

Tabelul poate fi gândit, de asemenea, ca reprezentând distribuția bidimensională a două variabile: *sexul* și *fumatul*. Dintr-un eșantion, de volum 107, s-au determinat pentru fiecare observație valorile celor două variabile:

observația	sexul	fumează sau nu
...	...	...
i	m	da
...	...	...

Datele au fost apoi sintetizate sub forma tabelului 7.17. În această situație prezintă interes studierea asocierii dintre cele două variabile. Cu alte cuvinte, este important să știm dacă, în cazul cunoașterii valorii unei variabile (pentru o observație), ni se oferă indicații asupra valorii celeilalte variabile (pentru aceeași observație). Aceste informații au aspect probabilist, adică, de exemplu: *dacă o persoană este bărbat, atunci există o șansă mai mare ca persoana să fumeze.*

În funcție de cele două moduri de interpretare a datelor, ipoteza nulă este:

- H_0 : repartiția variabilei *fumat* este identică în cele două populații;
- H_1 : nu există asociere între variabilele *sex* și *fumat*.

Indiferent de ipoteza nulă utilizată, distincția este una de interpretare; se va considera că situația observată se datorează sondajului (formării aleatorii a eșantionului), deci o configurație care se poate abate, în anumite limite rezonabile, de la o situație de referință proprie ipotezei nule. Testul va decide dacă „depărtarea" dintre situația de referință „așteptată" și situația observată poate fi atribuită hazardului sau este atât de mare încât poate fi atribuită mai degrabă adevărului ipotezei alternative (contrara ipotezei nule considerate).

Situația așteptată, în condițiile ipotezei nule, este arătată în tabelul 7.18, calculat potrivit principiului explicat la partea teoretică:

Tabelul 7.18. *Datele așteptate în exemplul 7.14*

	bărbați	femei	total
fumători	20,5	30,5	51
nefumători	22,5	33,5	56
total	43	64	107

De exemplu, 20,5 = 51×43/107. Se poate remarca, pentru un tabel 2×2, că numărul gradelor de libertate este egal cu 1, deoarece cunoașterea unei valori, de

exemplu prima casetă 20,5, conduce la fixarea celorlalte casete, datorită totalurilor care se păstrează. Astfel, 30,5 = 51–20,5; 22,5 = 43–20,5; 33,5 = 56–22,5 = 64–30,5.

Statistica testului este:

$$\chi^2 = \frac{(25-20{,}5)^2}{20{,}5} + \frac{(26-30{,}5)^2}{30{,}5} + \frac{(18-22{,}5)^2}{22{,}5} + \frac{(38-33{,}5)^2}{33{,}5} = 3{,}16.$$

Cum χ^2 calculat are un singur grad de libertate, $\nu = (2-1)\times(2-1) = 1$, după formula generală sau după discuția anterioară, iar din tabelul repartiției se obține $\chi^2_{1;0{,}95} = 3{,}84$, rezultă că valoarea calculată este mai mică decât valoarea critică a testului și, prin urmare, nu putem respinge ipoteza nulă în favoarea ipotezei alternative. Cu alte cuvinte, datele de sondaj nu susțin ipoteza că repartiția fumătorilor este diferită în cele două populații sau că fumatul este o practică influențată de sexul persoanei respective.

În cazul în care numărul total de observații este mai mic decât 75, valoarea χ^2 calculată tinde să supraestimeze asocierea celor două variabile. Din acest motiv, se recomandă utilizarea corecțiilor de continuitate (Yates), statistica testului devenind astfel:

$$\chi^2 = \sum_{i,j} \frac{(|O_{ij} - E_{ij}| - 0{,}5)^2}{E_{ij}}.$$

Exemplul 7.14, recalculat, produce o valoare de 2,49. Ipoteza nulă nu poate fi respinsă nici cu noua valoare.

Utilizarea corecției de continuitate este indicată pentru a contracara variabilitatea mărită a eșantioanelor mici.

Condiția generală de aplicabilitate a testului χ^2 este pentru tabelele 2×2 următoarea: toate cele patru frecvențe așteptate trebuie să fie egale sau mai mari decât 5.

Exemplul 7.15 (model k×2). Compararea a două populații după categoriile unei variabile nominale poate fi judecată și în termenii asocierii dintre variabile, dar acest mod de interpretare este mai puțin utilizat. Ipotezele și formulele de calcul sunt cele generale.

Într-un studiu privind studenții de la două profiluri (tehnic și economic), s-au înregistrat și datele despre categoria socioprofesională a părinților. Datele sunt prezentate sub formă de frecvențe absolute în tabelul 7.19. Se verifică ipoteza că distribuția categoriilor socioprofesionale ale părinților este aceeași indiferent de profilul facultății.

Tabelul este completat cu valorile așteptate. Acestea sunt trecute între paranteze în fiecare casetă. Sub această formă este mai ușor să se opereze cu datele respective. De menționat că, pentru simplificarea exemplului, s-au luat doar 5 clase ale scalei categoriilor socioprofesionale. Situația ar putea fi inversată prin interpretarea după care se compară cinci populații (determinate de categoria socioprofesională) după o variabilă cu două categorii nominale (profilul).

Tabelul 7.19. *Datele de sondaj ale exemplului 7.15*

	profil tehnic	profil economic	total
agricultori	8 (4,09)	2 (5,91)	10
muncitori	16 (12,28)	14 (17,72)	30
cadre medii	15 (17,19)	27 (24,81)	42
cadre superioare	26 (29,47)	46 (42,53)	72
alte categorii	14 (15,96)	25 (23,04)	39
total	79	114	193

Statistica χ^2 este

$$\chi^2 = \frac{(8-4{,}09)^2}{4{,}09} + \frac{(2-5{,}91)^2}{5{,}91} + \ldots + \frac{(25-23{,}04)^2}{23{,}04} = 9{,}79,$$

care provine, în ipoteza nulă, dintr-o repartiție χ^2 cu $\nu = (5-1)\times(2-1) = 4$ grade de libertate. La nivelul de semnificație $\alpha = 0{,}05$ se obține, din tabelul repartiției χ^2, quantila $\chi^2_{4;0{,}95} = 9{,}49$. Cum valoarea calculată este mai mare decât valoarea critică, se respinge ipoteza nulă a identității repartițiilor în cele două populații. Se poate accepta ipoteza că alegerea profilului este influențată de categoria socio-profesională a părinților.

Exemplul 7.16 (modelul general). Să considerăm la exemplul precedent un set mai complex de date, prin includerea mai multor profiluri (vezi tabelul 7.20). În acest caz, aplicarea unui test χ^2 este considerată, de obicei, o metodă de testare a asocierii dintre variabilele *profil* și *categorie socioprofesională.*

Tabelul 7.20. *Datele de sondaj pentru exemplul 7.16*

profilul	agricultori	muncitori	cadre medii	cadre superioare	alte categorii	total
tehnic	8 (5,11)	16 (11,5)	15 (18,15)	26 (28,38)	14 (15,85)	79
economic	2 (7,38)	14 (16,60)	27 (26,19)	46 (40,95)	25 (22,87)	114
umanist	6 (5,63)	10 (12,67)	22 (19,99)	31 (31,25)	18 (17,46)	87
artistic	4 (1,88)	5 (4,22)	7 (6,66)	8 (10,42)	5 (5,82)	29
total	20	45	71	111	62	309

Compararea globală a profilurilor este o procedură mai puțin intuitivă, întrucât respingerea ipotezei identității repartițiilor nu arată care dintre populațiile implicate sunt atât de diferite încât se respinge ipoteza nulă. În asemenea cazuri, după o comparare globală, se vor efectua teste de comparare pentru perechi de variabile, pentru a decide care pereche produce respingerea ipotezei nule.

Ipotezele și formulele de calcul sunt cele prezentate, singura precizare fiind aceea că, în cazul general, numărul gradelor de libertate este influențat de ambele dimensiuni (se aplică, evident, tot formula generală).

Se obţine $\chi^2 = 13{,}57$. Cum pentru $\nu = (4-1) \times (5-1) = 12$ grade de libertate (4 şi 5 numărul liniilor şi, respectiv, al coloanelor, fără totaluri) şi $\alpha = 0{,}05$ se obţine valoarea critică $\chi^2_{12\,0{,}95} = 21{,}03$, rezultă că ipoteza nulă, a lipsei de asociere între profil şi categorie socioprofesională, nu poate fi respinsă.

Este de remarcat că, spre deosebire de exemplul 7.15, concluzia este alta. Acest fapt se explică prin aceea că, datorită considerării unui număr mai mare de observaţii şi subpopulaţii, nu se mai susţine lipsa de omogenitate a datelor. Trebuie remarcat că într-o comparaţie globală nu contează situaţia relativă a eşantioanelor, unul faţă de celălalt, ci situaţia relativă a unui eşantion în raport cu ansamblul celorlalte. Din acest motiv se va efectua mai întâi o comparaţie globală şi, doar în cazul respingerii ipotezei nule, se vor efectua teste de comparare individuale.

7.6.14. Utilizarea testelor χ^2 în psihologie

În psihologie, testul χ^2 se foloseşte mai ales pentru următoarele tipuri de comparaţii:

- **o repartiţie observată şi o repartiţie teoretică.** De exemplu, compararea repartiţiei bărbaţilor şi femeilor, deputaţi în Parlamentul României în anul 1998, faţă de repartiţia conformă cu teoria egalităţii şanselor (50% – 50%). Ca rezultat al analizei statistice, se poate stabili dacă repartiţia deputaţilor după sex respectă teoria egalităţii şanselor;
- **mai multe repartiţii observate rezultate din mai multe eşantioane.** De exemplu, repartiţia bărbaţi/femei, deputaţi în Parlamentul României, după alegerile din 1992 şi 1996. Prin comparare, se poate spune dacă cele două repartiţii sunt semnificativ distanţate una de alta (reflectând o modificare în optica de completare a listelor de candidaţi de către partidele politice) sau nu. Este evident că rezultatul ar fi interpretat altfel, dacă alegerile parlamentare s-ar fi desfăşurat după sistemul uninominal.
- **mai multe repartiţii observate rezultate dintr-un singur eşantion, după diferitele modalităţi ale unei variabile calitative.** În acest caz, este posibilă studierea legăturii sau independenţei variabilelor calitative. De exemplu, pentru un eşantion de subiecţi înscrişi în partide politice, se poate studia legătura sau independenţa între variabila „apartenenţa la un partid politic dat" (variabilă calitativă cu două clase sau două modalităţi: dreapta/stânga) şi variabila „sex" (variabilă calitativă cu două clase). În urma analizei, se stabileşte dacă alegerea unei mişcări politice este asociată sexului.

Exemplul 7.17. Un grup de psihiatri a observat că, din 1900 de pacienţi cu sindroame specifice, 780 sunt muncitori, 523 sunt funcţionari cu studii medii, 380 sunt cadre superioare sau au profesiuni liberale şi restul aparţin altor categorii profesionale. Statisticile oficiale arată că, pentru aceleaşi categorii profesionale, repartiţia procentajelor este de 40%, 28%, 20% şi, respectiv, 12%. Se doreşte să se ştie dacă repartiţia după categoriile profesionale a lotului examinat o reflectă pe cea existentă la nivel naţional. Interpretarea rezultatului este:

- maladiile psihiatrice nu apar cu predilecţie la una sau mai multe categorii profesionale (cazul în care cele două repartiţii – repartiţia lotului şi repartiţia dată de statistica oficială – coincid sau nu sunt suficient de distanţate);
- anumite categorii profesionale sunt predispuse mai mult la maladii psihiatrice (cazul în care cele două repartiţii – repartiţia lotului şi repartiţia dată de statistica oficială – nu coincid, sunt suficient de distanţate).

Putem întocmi următorul tabel:

	Muncitori	Funcţionari	Cadre superioare	Alte categorii	Total
Efective observate	780	523	380	217	1900
Proporţii teoretice	40%	28%	20%	12%	100%
Efective teoretice	760	532	380	228	1900

Ultima linie, a efectivelor teoretice aşteptate, se obţine după procedura uzuală de înmulţire a proporţiilor teoretice cu efectivul total. Astfel, $760 = 40 \times 1900/100$.

Aprofundarea analizei χ^2

Din testul χ^2 de ajustare se poate trage o concluzie globală cuprinzând repartiţia efectivelor fiecărei categorii a unei variabile. Ca atare, acest test îşi atinge destul de repede limitele atunci când numărul de categorii devine important. Într-un astfel de caz, găsirea unei diferenţe semnificative este interpretată la nivel global, neindicându-se care categorie anume dă diferenţa. Pentru aflarea acestora, va fi necesară o analiză suplimentară, procedând la regrupări sau la analize de două câte două a efectivelor de modalităţi ale variabilelor.

M.B. Harris imaginează un dispozitiv experimental, relatat în articolul „The Effect of Performing one Altruistic Act on the Likelihood of Performing Another", publicat în *Journal of Social Psychology*, 88, 1972[51], şi prezentat în continuare.

În cadrul lucrărilor practice de psihologie socială, studenţii au decis să reia o experienţă care pune în aplicare o tehnică de manipulare bine cunoscută în literatura psihosocială. Obiectivul era obţinerea câtorva monede de la trecători. După caz, cererea monedelor era formulată diferit. Într-o primă condiţie, cererea era formulată direct („Scuzaţi, domnule, nu aveţi o monedă mică?"). Pentru alte două situaţii, aceeaşi cerere de bani era precedată de un act preparator, constând în a întreba mai întâi trecătorul abordat „Cât este ora?" (condiţia 2) sau în a cere o informaţie privind direcţia de urmat pentru a găsi o stradă din apropierea locului unde se efectua experimentul (condiţia 3). 50 de studenţi (băieţi între 18 şi 21 de ani) au participat la această experienţă. Fiecare testa cele trei condiţii experimentale. O

51. *Apud* Nicolas Gueguen, *Manuel de statistiques pour psychologues*, Dunod, Paris, 1997, pp. 188-200.

persoană (doar bărbaţi de vârstă matură) era testată pentru fiecare dintre condiţii. Cererea adresată subiecţilor testaţi a fost în prealabil repetată în cursul lucrărilor dirijate. Ordinea de repartiţie a condiţiilor era aleatorie. Rezultatele corespund numărului de persoane care au dat efectiv bani şi sunt prezentate pentru fiecare dintre condiţii.

Condiţia 1 Cerere directă	Condiţia 2 Întrebare oră	Condiţia 3 Întrebare stradă
10	23	21

Studenţii s-au întrebat dacă, la riscul de 5%, există o diferenţă între diferitele condiţii testate.

Au fost chestionate 150 de persoane şi, dintre acestea, un număr de 54 (10 +23+21) au acceptat să dea o monedă. Doar distribuţia acestora este conţinută în tabel. Sub ipoteza de echivalenţă strictă a condiţiilor, fiecare dintre casetele tabelului trebuie să aibă un efectiv teoretic egal cu 18 (=54/3). Rămâne să se măsoare abaterea între distribuţia observată (10, 23, 21) şi distribuţia teoretică (18, 18, 18). Valoarea χ^2 rezultă după următorul calcul:

Condiţia 1 Cerere directă	Condiţia 2 Întrebare oră	Condiţia 3 Întrebare stradă
$\frac{(10-18)^2}{18}=3,56$	$\frac{(23-18)^2}{18}=1,39$	$\frac{(21-18)^2}{18}=0,50$

Prin însumarea valorilor de pe ultima linie, se obţine $\chi^2 = 5,45$. Cu un număr de grade de libertate egal cu 2 (3 – 1), tabelul repartiţiei χ^2, la pragul de 5%, dă 5,99. Deoarece χ^2 calculat este inferior valorii critice, trebuie să tolerăm ipoteza nulă şi să admitem că diferenţele, observate între cele trei condiţii de efectuare a cercetării, rezultă din fluctuaţii de eşantionaj. Totuşi, dezechilibări între efective, după condiţiile manipulate, există în mod real. În condiţiile 2 şi 3, studenţii obţin „randamente" echivalente care par, din contra, să se distingă de cel obţinut în condiţia 1. În fapt, condiţiile 2 şi 3 sunt apropiate. Doar prima cerere este diferită, întrebarea despre oră sau despre stradă constituind act preparator pentru cererea finală. Se poate să ne întrebăm dacă rezultatul la test va fi acelaşi, în cazul în care efectivele condiţiilor 2 şi 3 sunt reunite. Testul se va efectua din acest moment asupra tabloului de repartiţie având următoarele efective:

Cerere directă	Act preparator
10	44

De această dată, distribuţia teoretică va consta într-un efectiv de 27 (54/2) în fiecare dintre căsuţe. Diferenţele dintre aceste două distribuţii (10, 44 contra 27 şi 27) vor da atunci rezultatul următor:

Cerere directă	Act preparator
$\frac{(10-27)^2}{27}=10,70$	$\frac{(44-27)^2}{27}=10,70$

Valoarea χ^2 obţinută este de 21,44, pe când valoarea critică (la un singur grad de libertate) este de 3,84. Diferenţa este de această dată semnificativă şi ne permite respingerea ipotezei de nul. Se poate spune că utilizarea unui act preparator permite accesul mai favorabil la satisfacerea unei anumite cereri.

Se remarcă faptul că cineva poate ajunge la distingerea randamentului între demersul direct şi cererea după un act preparator comparând două câte două dintre căsuţele tabelului iniţial. Rezultatele acestor comparaţii sunt prezentate în tabelul următor:

Comparaţia efectuată	Efective	χ^2 calculat	Concluziile $\alpha = 5\%$
Condiţia 1/Condiţia 2	10/23	5,12	Respingere H_0
Condiţia 1/Condiţia 3	10/21	3,90	Respingere H_0
Condiţia 2/Condiţia 3	23/21	0,09	Tolerare H_0

Se observă că situaţia cererii directe se distinge de cele două condiţii de manipulare, în timp ce ultimele două nu sunt distanţate semnificativ. Comparaţia generală efectuată iniţial nu permite evidenţierea unei diferenţe globale în repartiţie în funcţie de echivalenţa dintre condiţiile 2 şi 3. Avantajul unei comparări generale cu ajutorul unui aceluiaşi test conduce, în schimb, la creşterea dificultăţii de reperare a particularităţilor repartiţiilor observate. În majoritatea cazurilor trebuie să se rafineze analiza pentru a repera mai bine modalităţile sau grupurile de modalităţi care se disting sau nu se disting de alte modalităţi sau de alte grupuri de modalităţi. Dificultatea nu mai stă în calcul, ci în selecţia pertinentă a metodelor de analiză şi a elementelor de analizat.

7.7. Testarea normalităţii

Foarte multe dintre procedurile statistice presupun că datele de sondaj provin dintr-o repartiţie normală sau dintr-o repartiţie care nu se abate foarte mult de la cea normală. Verificarea acestei ipoteze este astfel esenţială şi trebuie să preceadă aplicarea unor prelucrări statistice bazate pe ipoteza normalităţii datelor.

Un prim procedeu poate fi bazat pe testele de concordanţă deja prezentate. De exemplu, printr-un test χ^2 de ajustare, se poate compara repartiţia de frecvenţe empirice cu repartiţia de frecvenţe teoretice (dedusă din ipoteza normalităţii variabilei studiate sau din orice altă repartiţie aşteptată). Aceeaşi verificare se realizează prin testul Kolmogorov-Smirnov.

Sunt însă construite şi teste specializate pentru verificarea normalităţii; unele dintre ele sunt prezentate în continuare.

7.7.1. Testul Lilliefors

Şi acest test verifică ipoteza nulă conform căreia datele observate provin dintr-un eşantion extras dintr-o populaţie normală faţă de ipoteza alternativă conform căreia repartiţia caracteristicii nu este normală. Aplicarea testului necesită efectuarea următoarelor etape:

- se calculează transformatele normate:

$$z_i = (x_i - \bar{x})/s \ , \ i = 1,2,\ldots,n,$$

unde $\bar{x}, s$ reprezintă media şi abaterea standard de sondaj.

- se ordonează valorile transformatelor z:

$$z_{(1)} \le z_{(2)} \le \ldots \le z_{(n)};$$

- se calculează funcţia de repartiţie de sondaj:

$$F_n(z_{(i)}) = \begin{cases} 0 & \text{pentru} \quad i \le 0 \\ \dfrac{1}{n} & \text{pentru} \quad 1 \le i \le n-1; \\ 1 & \text{pentru} \quad i = n \end{cases}$$

- se determină valorile corespunzătoare aşteptate, potrivit repartiţiei normale, $F(z_{(i)})$;
- se determină diferenţele absolute dintre valorile funcţiei de repartiţie de sondaj şi valorile aşteptate, $d_i = \left| F_n(z_{(i)}) - F(z_{(i)}) \right|$;
- se alege valoarea maximă a diferenţelor absolute, fie aceasta d_{max}. Valoarea obţinută este statistica testului Lilliefors.

Ipoteza nulă nu se respinge dacă $d_{\max} \le d_{n;1-\alpha}$, unde $d_{n;1-\alpha}$ este valoarea critică dată de tabelul 7.21, potrivit pragului de semnificaţie α ales.

Tabelul 7.21. *Valorile critice ale testului Lilliefors*

n	$1-\alpha$=0,95	$1-\alpha$=0,99	n	$1-\alpha$=0,95	$1-\alpha$=0,99
4	0,381	0,417	14	0,227	0,261
5	0,337	0,405	15	0,220	0,257
6	0,319	0,364	16	0,213	0,250
7	0,300	0,348	17	0,206	0,245
8	0,285	0,331	18	0,200	0,239
9	0,271	0,311	19	0,195	0,235
10	0,258	0,294	20	0,190	0,231
11	0,249	0,284	25	0,173	0,200
12	0,242	0,275	30	0,161	0,187
13	0,234	0,268	n>30	$\frac{0,886}{\sqrt{n}}$	$\frac{1,031}{\sqrt{n}}$

7.7.2. Teste bazate pe coeficienţii beta

Gradul de asimetrie sau boltirea graficului unei repartiţii de sondaj se poate aprecia prin coeficienţii β_1 şi β_2. Aceşti coeficienţi au, pentru repartiţia normală, valorile $\beta_1 = 0$ şi $\beta_2 = 3$. Se poate aprecia atunci depărtarea de la repartiţia normală prin testarea diferenţelor dintre valorile coeficienţilor de asimetrie şi de boltire de sondaj şi valorile menţionate pentru repartiţia normală.

Dacă notăm cu β_1^*, β_2^* valorile estimate din eşantion, se respinge ipoteza nulă a normalităţii datelor atunci când:

$$|Z_1| \approx \sqrt{\frac{n\beta_1^*}{6}} \ge z_{1-\alpha/2}$$

şi

$$|Z_2| \approx \left| (\beta_2^* - 3)\sqrt{\frac{n}{24}} \right| \ge z_{1-\alpha/2}.$$

Prima relaţie conduce, de fapt, la respingerea ipotezei că eşantionul provine dintr-o repartiţie simetrică (adică $\beta_1 = 0$), iar a doua relaţie conduce, de fapt, la respingerea ipotezei că eşantionul provine dintr-o repartiţie caracterizată de $\beta_2 = 3$. Respingerea ambelor ipoteze (sau doar a uneia dintre ele) poate arăta o neconcordanţă a distribuţiei empirice cu repartiţia normală.

Observaţie. În calcularea statisticilor testelor s-au utilizat formulele aproximative (există şi formule exacte, dar, pentru situaţiile frecvent întâlnite, este satisfăcător şi rezultatul obţinut prin formulele aproximative).

Capitolul 8

ASOCIEREA DATELOR

8.1. Generalități

Preocuparea pentru asocierea variabilelor apare o dată cu începuturile psihologiei experimentale și ale testului psihologic. Fundamentarea teoriei asocierii datelor în psihologie (și nu numai) aparține celebrei școli psihometrice engleze.

În 1888, F. Galton considera că legea deviației de la medie poate fi utilizată în scopul de a exprima gradientul relativ între două variabile. Cu aceasta introduce tehnica de corelație în biologie și psihologie. El propune un indicator simplu de apreciere a corelației, denumit indice de corelație. Lucrarea lui Galton, *Inquiries into Human Faculty and its Development*, dedicată diferențelor individuale în privința facultăților mentale, este considerată începutul mișcării testelor mentale.

Edge Worth creează noțiunea de coeficient de corelație, iar Karl Pearson, un elev al lui Galton, dezvoltă, prin utilizarea datelor cuprinse în încercările lui Bravais, forma finală a coeficientului de corelație prin momentul produselor. Galton, în colaborare cu Weldon, fondează în anul 1901 revista *Biometrica*, destinată cercetărilor în biologie și psihologie. În anul 1904, F. Galton a fondat și Laboratorul de biometrie al Universității din Londra, unde înființează o catedră de cercetare rezervată pentru K. Pearson.

În anul 1904, apar două lucrări de statistică aparținând lui Charles Spearman, după ce Clark Wissler, colaboratorul lui John McKeen Cattel, a prezentat, în anul 1901, lucrări de corelație statistică asupra testelor. Într-o primă lucrare, Spearman demonstrează importanța tehnicilor de corelație statistică pentru scopuri practice de cercetare, iar în a doua fundamentează tehnicile analizei factoriale, care încununează statistica proprie corelației, fapt ce prezintă mare importanță pentru cercetarea testelor de inteligență generală și a funcțiilor specifice. Spearman schițează totodată un model al ierarhiei inteligențelor. În 1901, împreună cu Krueger, studiază corelațiile între diverse randamente la testele de aptitudini.

Spunem, în general, că două (sau mai multe) variabile sunt asociate dacă, în distribuția comună a variabilelor, anumite grupuri de valori au probabilități mai mari de realizare simultană decât alte grupuri de valori. Cu alte cuvinte, dacă o variabilă ia o valoare, atunci celelalte variabile vor lua, cu probabilități mai mari, valori determinate de valoarea primei variabile.

Sensul corect al definiției este acela de consonanță, și nu de cauzalitate strictă. Definiția dată poate conduce ușor, din punct de vedere intuitiv, la ideea că prima variabilă determină, în sensul cauzalității, comportarea celorlalte variabile. Această interpretare este însă parțial eronată. Este adevărat că asocierile de tip cauză – efect se conformează definiției date, dar situațiile din lumea reală sunt mult mai diversificate.

Să presupunem că studiem relația dintre *înălțime* și *greutate* (folosim termenul „greutate" datorită uzitării sale în exprimarea comună; cei cunoscători vor identifica totuși noțiunea de masă) pe o populație umană. De regulă, persoanele cu înălțimi mai mari au și o greutate mai mare. Apar astfel asocieri de valori de tipul *înălțimi mari – greutăți mari* (și, prin compensare, *înălțimi mici – greutăți mici*). Vom spune atunci că variabilele înălțime și greutate sunt asociate. Prin urmare, pentru o persoană mai înaltă este mai probabil să măsurăm și o greutate mai mare. De aici însă până la a spune că înălțimea constituie cauza greutății este o distanță foarte mare.

Este evident că asociațiile ce reprezintă relații cauză – efect sunt importante în multe domenii de aplicații (să ne gândim, de exemplu, la medicină), dar demonstrarea cauzalității nu este nici simplă, nici directă și nici de necontestat. Pentru exemplificare, redăm **criteriile lui Hill** de evaluare a plauzibilității unei relații cauză – efect:

- planul de studiu după care s-a evidențiat asocierea (modul de obținere a datelor experimentale, studiul lor etc.);
- gradul asocierii – cu considerarea semnificației statistice;
- consistența – aceeași asociere a fost sesizată de mai mulți investigatori, prin diferite modele de studiu, în locuri și circumstanțe diferite ;
- relația temporală (cauza trebuie să preceadă efectul);
- relația doză – răspuns (este de presupus că o mai mare „cantitate de cauză" trebuie să producă un efect mai mare);
- plauzibilitatea – o relație trebuie să fie plauzibilă, adică trebuie să fie corelată cu nivelul de cunoaștere curentă disponibilă (cerința nu este, în sine, nici necesară, nici suficientă);
- specificitatea – dacă o singură cauză poate fi legată de un singur efect, această specificitate furnizează o evidență directă în favoarea cauzalității;
- analogia – existența unor relații cauzale analoage poate oferi condiții pentru o nouă cauzalitate.

Se observă că subiectul cauzalității relevate de studiile experimentale impune un studiu multiplu și remarcăm că primele două criterii (și nu numai ele) sunt legate strict de statistică. În același timp, se poate spune că statistica singură nu poate evidenția legături cauzale în lumea reală. Reținem însă că un aspect al problemei este **evidențierea asocierii** și alt aspect este **evaluarea gradului de asociere.**

Din punctul de vedere al utilității, evidențierea asocierii dintre variabile poate conduce la simplificarea unor proceduri de analiză a realității: faptul că o variabilă

este asociată suficient de puternic cu altă variabilă poate duce la abandonarea determinării experimentale pentru una dintre variabile (de regulă, cea care se obține cu cheltuieli mai mari de resurse). Cu alte cuvinte, o variabilă conține aproximativ aceeași informație, care este oferită de variabila asociată; se poate reduce investigarea doar la una dintre cele două variabile.

Principiul parcimoniei este, de exemplu, esențial în analiza factorială, care își propune identificarea unui număr redus de variabile subiacente ce pot să caracterizeze un ansamblu complex de date observate.

Un alt aspect este acela în care cunoașterea asocierii dintre variabile permite îmbunătățirea prognozelor efectuate pentru una dintre variabile dacă se cunoaște valoarea celeilalte.

Se observă astfel că, în analiză, una dintre variabile joacă rolul de **variabilă dependentă**, iar cealaltă (sau celelalte) joacă rolul de **variabilă independentă (variabile independente).** Denumirile (ca și rolurile) sunt relative, doar necesitățile și posibilitățile cercetătorului stabilind rolul fiecărei variabile. Acceptăm totuși că în studiul asocierii suntem interesați să cunoaștem ce se întâmplă cu valorile unei variabile (variabila dependentă) cunoscând valorile celorlalte variabile (variabilele independente).

***

Gradul de asociere este exprimat prin coeficienți care, pe cât posibil, sunt standardizați (adimensionali) pentru a permite comparații directe. Există mai multe scale pe care se măsoară gradul de asociere. Se disting scale standardizate direcționale și scale standardizate nedirecționale.

Scala direcțională (–1...+1). Permite distingerea asocierilor negative (în care valorile variabilelor sunt invers proporționale: mare cu mic, mic cu mare) de asocierile pozitive (în care valorile variabilelor sunt direct proporționale: mare cu mare, mic cu mic). Valoarea zero indică lipsa de asociere. Este specifică variabilelor ordinale și celor continue (de interval).

–1	0	+1
Asociere negativă (inversă)	Lipsă de asociere	Asociere pozitivă (directă)

Scala nedirecțională (0...1). Este mai puțin sensibilă decât scala direcțională, permite doar diferențierea asociere – neasociere, fără a specifica sensul asocierii. Pentru variabilele asociate puternic nu se poate preciza tipul de asociere. Este specifică variabilelor nominale.

Pentru un indicator de asociere se poate defini **atributul PRE** (Proportional Reduction in Error) – reducerea proporțională a erorii. Acest atribut se referă la îmbunătățirea prognozei privind valorile variabilei dependente atunci când se cunoaște valoarea variabilei independente. Nu toți indicatorii de asociere au atributul PRE, dar cei care îl au sunt preferați celor care sunt lipsiți de așa ceva.

Să considerăm un studiu în care 100 de salariați au răspuns la două întrebări: cum sunt plătiți și cât sunt de mulțumiți de munca lor. Răspunsurile la întrebări (valori de variabile dihotomice, nominale) sunt prezentate în următorul tabel de contingență:

		Plata		
		Bună	Rea	Total
Satisfacția muncii	Satisfăcut	35	15	50
	Nesatisfăcut	15	35	50
	Total	50	50	100

Considerăm că „satisfacția muncii" este variabila dependentă, în timp ce „plata" este variabila independentă (manipularea variabilei „plata" poate duce la modificarea variabilei „satisfacția muncii"[52]). Dorim deci să putem prognoza (prezice, evalua) gradul de mulțumire a unei persoane cunoscând modul în care este plătită. O analiză sumară a tabelului arată că există o anumită asociere între valori, deoarece repartizarea celor 100 de persoane nu este uniformă.

Analizând tabelul, putem spune că o predicție corectă poate fi făcută în 35 + 35 = 70 (%) din cazuri dacă știm care este nivelul de plată și utilizăm frecvențele maxime (plată bună => satisfăcut, plată rea => nesatisfăcut). Prin urmare, avem 30% erori.

Totuși, acest raționament dă prea multă putere de predicție modului de plată, deoarece, chiar în lipsa oricărei cunoașteri privind plata, putem greși, în orice predicție, doar în 50% dintre cazuri (cunoscând repartiția marginală 50 – 50 a satisfacției muncii).

Se definește atunci o măsură a cantității cu care cunoașterea variabilei independente mărește corectitudinea predicției în raport cu o „ghicire" pură (bazată doar pe cunoașterea repartiției variabilei dependente):

$$\mathrm{PRE} = \frac{\varepsilon_{dep} - \varepsilon_{ind}}{\varepsilon_{dep}},$$

unde ε_{dep} este eroarea folosind doar variabila dependentă, iar ε_{ind} este eroarea folosind variabila independentă.

În exemplul considerat, PRE = (50 – 30)/50 = 40(%), adică s-a redus cu 40% eroarea de predicție prin utilizarea informației date de variabila independentă. Se poate gândi PRE ca fiind varianța explicată de variabila independentă (deci 40% din varianța variabilei dependente este explicată de variabila independentă).

52. Situația prezentată este una simplificată; studiile psihosociologice arată că satisfacția muncii este multicondiționată (de altfel, este cazul cel mai frecvent în disciplinele psihosociale).

8.2. Asocierea între două variabile nominale (calitative)

Considerăm două variabile discrete nominale, fie ele X şi Y, având r şi, respectiv, s categorii. Studierea legăturii între cele două variabile are la bază un tabel de frecvenţe încrucişate, numit şi tabel de contingenţă. Termenul de contingenţă are sens asemănător celui de contingent, utilizat în armată, adică de grup omogen de indivizi în privinţa anumitor caracteristici.

Situaţia generală este prezentată în tabelul următor:

	Y_1	...	Y_j	...	Y_s	Total
X_1	n_{11}	...	n_{1j}	...	n_{1s}	$n_{1\bullet}$
...	...	...	...	...	...	...
X_i	n_{i1}	...	n_{ij}	...	n_{is}	$n_{i\bullet}$
...	...	...	...	...	...	...
X_r	n_{r1}	...	n_{rj}	...	n_{rs}	$n_{r\bullet}$
Total	$n_{\bullet 1}$	...	$n_{\bullet j}$	...	$n_{\bullet s}$	n

În acest tabel de contingenţă se pot distinge 4 tipuri de efective: efective încrucişate, efective condiţionale, efective marginale şi un efectiv total. Se notează cu n_{ij} efectivul încrucişat al cuplului (i, j), reprezentând numărul de observaţii care corespund modalităţii i a variabilei X şi modalităţii j a variabilei Y. Efectivul total este $n = \sum_i \sum_j n_{ij}$ şi reprezintă volumul eşantionului.

Efectivele marginale sunt calculate pe ultima linie şi pe ultima coloană şi sunt notate prin $n_{i\bullet}$ şi $n_{\bullet j}$, unde $n_{i\bullet}$ reprezintă totalul frecvenţelor de pe linia i, iar $n_{\bullet j}$ este suma frecvenţelor de pe coloana j. Introducerea în notaţii a caracterului (*) este necesară pentru a identifica frecvenţele marginale pe linie şi pe coloană, localizarea caracterului (*) fixând sensul în care s-a efectuat însumarea. De exemplu, $n_{i\bullet}$ semnifică însumarea elementelor de pe linia i (efectuată după toate coloanele acestei linii). Efectivele marginale nu sunt niciodată nule (totdeauna $n_i>0$ şi $n_j>0$), prin logica cercetării. Nu prezintă nici un interes considerarea unei modalităţi care nu este reală. Este însă evident că pot exista casete cu efectiv zero, dacă anumite intersecţii de modalităţi nu sunt întâlnite în cercetarea curentă. Asemenea situaţii trebuie însă evitate.

O primă analiză care se poate efectua asupra unui tabel de frecvenţe este trecerea la frecvenţe relative şi la comparaţii uzuale între acestea. Se pot obţine trei tipuri de frecvenţe relative:

- raportate la totalurile pe linii ($n_{ij}/n_{i\bullet}$), caz în care se obţine tabloul profilurilor-linii şi se pot compara liniile între ele: prin ponderarea cu totalurile liniilor, se pot compara subpopulaţiile determinate de nivelurile variabilei X;
- raportate la totalurile pe coloane ($n_{ij}/n_{\bullet j}$), caz în care se obţine tabloul profilurilor-coloane şi se pot compara coloanele între ele: prin ponderarea cu totalurile coloanelor, se pot compara subpopulaţiile determinate de nivelurile variabilei Y;
- raportate la n, volumul eşantionului. Se notează cu $f_{ij} = n_{ij}/n$ frecvenţa relativă încrucişată a cuplului (i, j), iar cu $f_{i\bullet}=n_{i\bullet}/n$ şi $f_{\bullet j}=n_{\bullet j}/n$ frecvenţele relative marginale. Rezultă că

$$\sum_i \sum_j f_{ij} = 1.$$

Frecvenţele relative obţinute prin raportarea la frecvenţele marginale se mai numesc şi frecvenţe relative condiţionale. Frecvenţa lui i condiţionată de j este raportul dintre efectivul încrucişat n_{ij} şi efectivul marginal $n_{i\bullet}$. Acest raport se notează f_j^i (se citeşte f de i, dacă j).

Familia (f_k^j) a k frecvenţe de k şi j se numeşte profilul lui j. Familia (f_k) a K frecvenţe marginale este numită profil mediu (sau marginal) pe K. La fel, frecvenţa lui j condiţionată de k sau frecvenţa de j dacă k este notată f_j^k. Astfel $f_j^k = \frac{n_{jk}}{n_k}$. De unde profilul lui k: f_j^k, familia a J frecvenţe de j dacă k şi profilul mediu asupra lui J:f_j, familia de J frecvenţe.

***

Se poate considera că variabilele X şi Y sunt independente empiric, dacă are loc relaţia

(1) $$n_{ij} = \frac{n_{i\bullet} \times n_{\bullet j}}{n}, (\forall) i, j,$$

care exprimă faptul că liniile (coloanele) sunt proporţionale între ele, deci profilurile-linii sunt egale (ca şi profilurile-coloane). Se poate demonstra că, în aceste condiţii, proporţia elementelor clasate Y_j este aceeaşi între elementele clasate X_i şi între elementele neclasate X_i (elementele non X_i). Această observaţie întăreşte ideea de independenţă între caracteristicile X_i şi Y_j, independenţă care, extinsă la toate nivelurile, conduce la independenţa între X şi Y.

Orice abatere de la egalităţile exprimate de relaţia (1) denotă o îndepărtare de independenţă a celor două variabile, adică poate constitui baza unui indicator al asocierii dintre variabile.

Vom spune că între nivelurile X_i şi Y_j:

- există o asociere pozitivă dacă $n_{ij} > (n_{i\bullet} n_{\bullet j}) / n$ (adică în caseta ij există o frecvenţă mai mare decât în populaţie);
- există o asociere negativă dacă $n_{ij} < (n_{i\bullet} n_{\bullet j}) / n$ (adică în caseta ij există o frecvenţă mai mică decât în populaţie).

Se adoptă ca măsură a asocierii (globale) mărimea (Pearson):

$$d^2 \text{ (sau } \chi^2) = \sum_{i,j} \frac{\left(n_{ij} - \frac{n_{i*} \times n_{*j}}{n} \right)^2}{\frac{n_{i*} \times n_{*j}}{n}}$$

sau $d^2 = n\left[\sum_{ij} \frac{n_{ij}^2}{n_{i*} n_{*j}} - 1 \right]$.

De notat că pentru fiecare căsuță cantitatea

$$\frac{\left(n_{ij} - \frac{n_{i*} \times n_{*j}}{n} \right)^2}{\frac{n_{i*} \times n_{*j}}{n}} \times \frac{1}{\chi^2},$$

numită contribuția la χ^2, evidențiază asociațiile semnificative.

Cum $n_{ij}/n_{*j} \le 1 \Rightarrow \frac{n_{ij}^2}{n_{*j} n_{ij}} \le \frac{n_{ij}}{n_{*j}} \Rightarrow d^2 \le n(s-1)$ și, analog, $d^2 \le n(r-1)$, de unde rezultă că

$$\frac{d^2}{n} \le \min(s-1, r-1),$$

adică există o limită superioară pentru cantitatea d^2/n, deși d^2 este dependentă de ordinul de mărime al frecvențelor (deci al eșantionului) și de numărul de niveluri ale variabilelor X și Y.

Marginea superioară este atinsă în cazul dependențelor funcționale (doar câte o căsuță nenulă pe fiecare linie, fiecare coloană).

Indicatorul χ^2 (numit indicator al contingenței pătratice) nu este măsurat pe o scală 0 –1 și, din acest motiv, neîncadrându-se în teoria generală a indicilor de asociere, nu poate fi utilizat cu bune rezultate pentru comparații.

Au fost propuși atunci o serie de alți indici (derivați din χ^2):

$\phi^2 = \frac{\chi^2}{n}$, indicator al contingenței pătratice medii (Pearson)

$$C = \sqrt{\frac{\chi^2}{n + \chi^2}},$$

coeficientul contingenței pătratice medii (Pearson), cu observațiile:

- ϕ^2 are neajunsul că dă rezultate diferite, pe aceeași populație, atunci când se modifică modul de grupare în clase;
- același lucru se poate spune și despre C. În plus, C nu poate lua valoarea maximă 1 (atinsă doar la limită, după dimensiunea tabloului). Pentru o asociere perfectă, la o repartiție simetrică de frecvențe r×r (atât X, cât și Y au r niveluri)

$$C = \sqrt{\frac{r-1}{r}}.$$

Utilizarea lui C este mai potrivită pentru repartiții de dimensiuni mai mari: la 5×5 C poate fi 0,80 etc.

Pornind tot de la χ^2, se definește coeficientul lui Ciuprov:

$$T^2 = \frac{\varphi^2}{\sqrt{(s-1)(t-1)}},$$

care, pentru tablouri pătrate, r = s, variază de la 0 la 1, deci se conformează unei măsuri de asociere nedirecțională.

În sfârșit, coeficientul lui Cramer:

$$V^2 = \sqrt{\frac{\phi^2}{\min\{(s-1),(r-1)\}}}.$$

Semnificația statistică a asocierii

Se poate utiliza cantitatea χ^2 pentru a testa asocierea evidențiată de datele de sondaj. Testul este unul de ajustare (concordanță) și reamintim că testul de asociere χ^2 are ipotezele:

- $\mathbf{H_0}$: variabilele X și Y nu sunt asociate;
- $\mathbf{H_1}$: variabilele X și Y sunt asociate.

Se demonstrează că, în condițiile ipotezei H_0, variabila χ^2, calculată mai sus, este repartizată χ^2 cu un număr de grade de libertate egal cu $(r-1)(s-1)$, $\chi^2 \sim \chi^2_{(r-1)(s-1)}$, ceea ce permite efectuarea unui test statistic.

Decizia în test se poate lua:

- pe baza comparării valorii calculate χ^2 cu valoarea critică tabelată corespunzătoare pragului de semnificație α fixat; se respinge ipoteza nulă dacă valoarea calculată este mai mare sau egală cu valoarea critică tabelată;
- pe baza interpretării probabilității critice a testului; ipoteza nulă se respinge dacă probabilitatea critică ce a fost calculată este mai mică decât pragul de semnificație α fixat.

În cazul în care nu se respinge ipoteza nulă, nu înseamnă că între cele două variabile nu există asociere. Trebuie interpretat că *asocierea nu a putut fi pusă în evidență de datele existente* sau că *asocierea nu este statistic semnificativă.*

Prin respingerea ipotezei nule se afirmă doar existența asocierii dintre variabile, nu și măsurarea gradului de asociere. Pentru aceasta se utilizează alte tehnici statistice, de exemplu coeficienții de corelație.

Limite de aplicare. Reamintim că pentru a aplica testul χ^2 trebuie ca nu mai mult de 20% dintre căsuțe să aibă mai puțin de 5 elemente.

Exemplul 8.1. O distribuție de efective (sau de frecvențe absolute) produsă de două variabile categorizate, indexată prin J și K, este numită tabel de contingență de efective (sau de frecvențe) pe J×K. Elementele j ale lui J și k ale lui K sunt numite categorii sau modalități ale variabilei.

Să avem în vedere cazul unei cercetări efectuate pe o populație școlară. Eșantionul cuprinde 955 de copii și s-au urmărit două variabile. Tabelul de contingență obținut prin încrucișarea celor două variabile, notate J și K, este prezentat în continuare. Variabila „studiile părinților" are 4 modalități ale nivelului de studii pe care le posedă tatăl copilului și „reușita școlară", notată K, cu trei niveluri de reușită școlară (slabă, medie, foarte bună).

	k1	**k2**	**k3**	**Total**
j1	17	356	40	413
j2	29	228	36	293
j3	54	108	25	187
j4	17	35	10	62
Total	117	727	111	955

Modalitățile variabilei J sunt j1 – studii generale, j2 – studii medii, j3 – studii universitare, j4 – studii postuniversitare. Pentru variabila K avem următoarele modalități: k1 – reușită școlară de nivel bun, k2 – reușită medie, k3 – reușită școlară slabă.

Pentru fiecare dintre cele 4 modalități ale variabilei J stabilim efective condiționale. Avem astfel 4 efective condiționale ale variabilei J și 3 efective condiționale ale variabilei K.

Distribuția condițională a copiilor cu tatăl având studii generale în funcție de rezultatele școlare este următoarea

	k1	k2	k3	Total
j1	17	356	40	413

Distribuția condițională a copiilor cu tatăl având studii medii în funcție de rezultatele școlare:

	k1	k2	k3	
j2	29	228	36	293

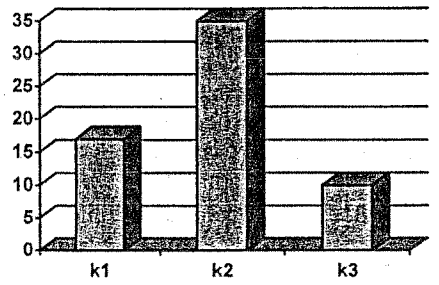
Distribuția condițională a copiilor având tatăl cu studii superioare, în funcție de reușita școlară, este următoarea:

	k1	k2	k3	
j3	54	108	25	187

Distribuția parțială a copiilor cu tatăl având studii postuniversitare, în funcție de reușita școlară:

	k1	k2	k3	
j4	17	35	10	62

Dacă prezintă importanță, pentru fiecare dintre aceste distribuții vom face grafice adecvate. Astfel, distribuția condițională a copiilor cu tatăl având studii postuniversitare, în funcție de reușita școlară, este următoarea:



De asemenea, mai pot exista următoarele distribuții condiționale:

Distribuția condițională a rezultatelor școlare de bun nivel în funcție de nivelul de studii al tatălui:

	k1
j1	17
j2	29
j3	54
j4	17

Distribuția condițională a rezultatelor școlare de nivel mediu în funcție de profesia tatălui:

	k2
j1	356
j2	228
j3	108
j4	35

Distribuția condițională a rezultatelor școlare de bun nivel în funcție de nivelul de studii al tatălui;

	k3
j1	40
j2	36
j3	25
j4	10

De asemenea, distribuțiile condiționale pot fi reprezentate grafic. În total avem deci 7 distribuții parțiale sau condiționale.

Distribuţiile marginale reflectă fidel efectivele unei variabile. Astfel, distribuţia marginală a variabilei „nivel de studii", notată de noi cu J, este următoarea:

j1	413
j2	293
j3	187
j4	62

Distribuţia parţială sau marginală a variabilei „reuşită", notată de noi prin K, este următoarea:

	k1	k2	k3	Total
j1	17	356	40	413

Desigur, situaţia poate fi prezentată şi sub forma unor frecvenţe relative.

Totalul general este de 955 de subiecţi. Suma frecvenţelor marginale este de 955, indiferent dacă însumarea are loc pe linii sau pe coloane. Acest rezultat poate constitui un criteriu pentru aprecierea corectitudinii calculelor.

După ce am prezentat frecvenţele observate într-o situaţie concretă, vom trece la stabilirea efectivelor teoretice, care se stabilesc în virtutea unei ipoteze pe care o lansăm, de obicei ipoteza de nul, sau prin comparare cu situaţia generală existentă. Sunt valabile aceleaşi precizări pe care le-am făcut la începutul tratării testului chi pătrat. Din punctul de vedere al tehnicii de calcul, fiecare frecvenţă încrucişată teoretică se obţine prin înmulţirea valorilor corespunzătoare, din linie şi coloană, ale frecvenţelor marginale şi prin împărţirea acestui produs la totalul general, egal cu 955.

Astfel, frecvenţa teoretică aflată la intersecţia primei linii şi primei coloane se obţine prin $F_{11} = \frac{243 * 103}{841} = 29{,}76$. Procedând similar, obţinem toate frecvenţele încrucişate teoretice. Prezentăm mai jos tabelele cu efectivele observate şi efectivele teoretice:

	k1	k2	k3	Total
j1	17	356	40	413
j2	29	228	36	293
j3	54	108	25	187
j4	17	35	10	62
Total	117	727	111	955

	k1	k2	k3	Total
j1	50,60	314,40	48,00	413
j2	35,90	223,06	34,06	293
j3	22,91	142,35	21,74	187
j4	7,60	42,20	7,21	62
Total	117	727	111	955

Urmând procedura de calcul al valorii χ^2, obţinem următorul tablou:

f_o	f_t	F_0-f_t	$(f_o-f_t)^2$	$(f_0-ft)^2/f_t$
17,00	50,60	–33,60	1128,96	22,31
356,00	314,40	41,60	1730,56	5,50
40,00	48,00	–8,00	64,00	1,33
29,00	35,90	–6,90	47,61	1,33
228,00	223,06	4,94	24,40	0,11
35,00	34,06	0,94	0,88	0,03
54,00	22,91	31,09	966,59	42,19
108,00	142,35	–34,35	1179,92	8,29
25,00	21,74	3,26	10,63	0,49
17,00	7,60	9,40	88,36	11,63
35,00	42,20	–7,20	51,84	1,23
10,00	7,21	2,79	7,78	1,08
			χ^2	95,51

Am obţinut $\chi^2 = 95{,}51$. Pentru $(4 - 1)(3 - 1) = 6$ grade de libertate, se respinge ipoteza nulă ($p < 0{,}01$).

Datele pentru tabelul χ^2 mai pot fi prezentate şi în felul următor:

	k1	k2	k3
J1	–33,60	+41,60	–8,00
J2	–6,90	+4,94	+ 0,94
J3	+ 31,09	–34,35	+3,26
J4	+ 9,40	–7,20	+2,79

Sau pur şi simplu aşa:

	k1	k2	K3
J1	–	+	+
J2	–	+	+
J3	+	–	+
J4	+	–	+

În tabelele anterioare se face diferenţa dintre frecvenţele observate şi frecvenţele teoretice. În primul tabel se trec valorile obţinute, însoţite de semnul corespunzător, iar în al doilea tabel numai semnul.

Dacă frecvenţa observată este mai mare decât frecvenţa teoretică, se spune că legătura dintre j şi k este pozitivă sau că există o „atracţie". Dacă frecvenţa absolută este mai mică decât frecvenţa teoretică, se spune că legătura dintre j şi k este negativă sau că există o „repulsie". Când sunt egale, se spune că există o legătură nulă.

8.3. Asocierea datelor ordinale

Datorită faptului că scalele de măsură ordinale au definită ordinea între categorii (clase), coeficienţii de asociere sunt direcţionali, adică măsuraţi pe o scală (–1...0...+1).

8.3.1. Coeficientul de corelaţie a rangurilor (Spearman) r_s

Acest coeficient reprezintă o măsură PRE şi este direcţional. Se utilizează atunci când observaţiile cuprind valori ale unor variabile ordinale, structurate ca în tabelul următor:

Observaţia	variabila 1	variabila 2
1	v_{11}	v_{21}
2	v_{12}	v_{22}
...	...	...
n	v_{1n}	v_{2n}

Se începe prin a atribui ranguri valorilor fiecărei variabile, separat. În atribuirea de ranguri se vor trata situaţiile de *ex aequo* prin atribuirea ca rang a unei secvenţe de valori egale a rangului obţinut ca media aritmetică a rangurilor cuvenite valorilor egale. Se obţine astfel:

Observaţia	ranguri variabila 1	ranguri variabila 2	d^2
1	r_{11}	r_{21}	$(r_{11}-r_{21})^2$
2	r_{12}	r_{22}	$(r_{12}-r_{22})^2$
...	...	...	...
n	r_{1n}	r_{2n}	$(r_{1n}-r_{2n})^2$

Formula de definiţie a coeficientului de corelaţie a rangurilor este:

$$r_s = 1 - \frac{6\sum_i d_i^2}{n(n-1)}.$$

Trebuie menţionat faptul că r_s este sensibil la *ex-aequo*-uri şi, în cazul unui mare număr de egalităţi, se recomandă utilizarea coeficientului de corelaţie liniară (prezentat la asocierea datelor continue), care produce o estimare mai bună a corelaţiei din populaţie.

Deoarece r_s este calculat pe baza datelor unui eşantion, rezultă că r_s este o statistică, reprezentând o estimaţie a coeficientului de corelaţie ρ_s existent la nivel de populaţie.

Testarea semnificaţiei coeficientului calculat r_s se bazează pe statistica t, dată de:

$$t = r_s\sqrt{\frac{n-2}{1-r_s^2}},$$

unde n este volumul eşantionului, iar r_s este coeficientul de corelaţie Spearman calculat. În ipoteza nulă, H_0: $\rho_s = 0$, statistica t urmează o repartiţie Student cu $\nu = n - 2$ grade de libertate. Se va respinge H_0 în favoarea unei ipoteze alternative, direcţionale sau nu, prin procedura uzuală a unui test bazat pe repartiţia Student. Atenţionăm că ipoteza verificată se referă la populaţie, şi nu la eşantion (este formulată pentru ρ_s, şi nu pentru r_s).

Există şi tabele de verificare directă a unui coeficient de corelaţie Spearman. Valoarea calculată este comparată în acest caz cu valoarea tabelată corespunzătoare volumului eşantionului şi pragului de semnificaţie fixat. Ipoteza nulă se respinge dacă valoarea calculată este mai mare decât valoarea tabelată.

Acest coeficient de corelaţie este neparametric şi se foloseşte mai ales atunci când numărul de subiecţi este mic (sub 50 de perechi de valori). Coeficientul de corelaţie a rangurilor nu este adecvat pentru scale ordinale cu puţine categorii (5-7) care dau, de regulă, multiple cazuri de egalitate: pe o scală cu 5 categorii, un eşantion de 30 de elemente duce, într-o distribuţie uniformă, la 6 egalităţi pe categorie, ceea ce înseamnă mult. În asemenea situaţii se utilizează coeficientul γ, definit în continuare.

Exemplul 8.2. Să presupunem că dispunem de rezultatele de la o probă de atenţie (AD-Praga) şi de la o probă de dexteritate manuală (O'Connor). Prin X vom nota rezultatele la proba de atenţie şi prin Y rezultatele la proba de dexteritate manuală. Datele observate sunt în tabelul următor:

Nr.crt.	X	Y	Rx	Ry	d	d^2
1	47	6	13,5	14,5	–1	1,00
2	52	9	10,0	12,0	–2	4,00
3	75	17	2,0	1,5	0,5	0,25
4	81	10	1,0	10,5	–9,5	90,25
5	66	11	5,5	8,5	–3,0	9,00
6	39	7	15,0	13,0	2	4,00
7	47	6	13,5	14,5	–1	1,00
8	55	11	10,0	8,5	1,5	2,25
9	61	17	8,0	1,5	6,5	42,25
10	70	13	3,0	5,0	–2	4,00
11	66	12	5,5	6,5	–1	1,00
12	49	14	12,0	4,0	8	64,00
13	59	10	9,0	10,5	–1,5	2,25
14	63	12	7,0	6,5	0,5	0,25
15	69	15	4,0	3,0	1	1,00
Total						226,75

Din aplicarea formulei rezultă:

$$r_s = 1 - \frac{6 \times 226{,}75}{15(15^2 - 1)} = 1 - \frac{1364{,}50}{3360} = 1 - 0{,}45 = 0{,}55.$$

În tabelul de corelații critice pentru testul bilateral se găsește 0,5214, pentru $\alpha = 0{,}05$. Deoarece valoarea calculată este mai mare decât valoarea critică, se poate respinge ipoteza nulă. Acceptăm, prin urmare, concluzia că în populație există o asociere semnificativă (fără stabilirea sensului de abatere de la zero) între variabilele studiate. Continuând cu testul unilateral, se poate respinge ipoteza nulă în favoarea ipotezei alternative care stipulează că în populație există o asociere pozitivă între cele două variabile. Astfel, se poate accepta că rezultate bune la proba de atenție conduc la rezultate bune și la proba de dexteritate manuală.

8.3.2. Coeficientul gama (γ)

Acest indicator se bazează tot pe atribuirea de ranguri valorilor celor două variabile ordinale, dar pleacă de la problema: *cunoașterea ordinii a două ranguri pentru variabila independentă poate prezice ordinea rangurilor pentru variabila dependentă*? Răspunsul ar trebui să fie afirmativ pentru variabile asociate, orice abatere de la acest răspuns trebuind să conducă la ideea de independență (lipsă de asociere) între cele două variabile.

Coeficientul γ este o măsură PRE pe o scală direcțională și se definește prin

$$\gamma = \frac{f_a - f_i}{f_a + f_i},$$

unde f_a este numărul de concordanțe (potriviri), f_i este numărul de inversiuni, spunând că

$$\begin{matrix} r_{1i} & r_{2i} \\ \dots & \dots \\ r_{1j} & r_{2j} \end{matrix}$$

se potrivesc dacă ordinea este aceeași în cele două coloane („$r_{1i} < r_{1j}$ și $r_{2i} < r_{2j}$" sau „$r_{1i} > r_{1j}$ și $r_{1i} > r_{1j}$") sau că are loc o inversiune dacă ordinea este schimbată ($r_{1i} < r_{1j}$ și $r_{2i} > r_{2j}$ sau relația inversă). De exemplu, în tabelul următor:

Observația	Rx	Ry
1	13,5	14,5
2	10,0	12,0
3	2,0	1,5
4	16	5

observația 1 se „potrivește" cu observația 3 (deoarece rangurile sunt în aceeași ordine în coloanele Rx și Ry: 13,5 > 2 și 14,5 > 1,5), iar observația 2 este inversată față de observația 4 (deoarece rangurile sunt în ordine inversă în coloanele Rx și Ry: 10 < 16 și 12 > 5).

Pentru calcule, se ordonează observațiile după prima variabilă (independentă) și apoi se numără inversiunile ca la permutările de n obiecte (din algebră). *Perechile care coincid în cel puțin o variabilă (nu sunt nici concordanțe, nici inversiuni) se exclud din calcule.*

În cazurile de multiple egalități, cum ar fi datele prezentate în tabelul următor:

consum de alcool / criminalitate	foarte înalt	înalt	scăzut	foarte scăzut
foarte înaltă	a	b	c	d
înaltă	e	f	g	h
scăzută	i	j	k	l
foarte scăzută	m	n	o	p

calculul potrivirilor este efectuat din colțul din stânga-sus, pentru fiecare căsuță adunând frecvențele din casetele dominate de căsuța respectivă (casetele așezate mai jos și mai la dreapta):

- caseta (a) domină blocul de căsuțe f : p;
- caseta g domină blocul l și p;
- caseta h nu domină nici o altă căsuță etc.

Se obțin astfel potrivirile:

- pentru (a): a · (f + g + h + j + k + l + n + o + p);
- pentru (b): b · (g + h + k + l + o + p);
- ...
- pentru (k): k · p.

Numărul total, f_a, se obține ca fiind suma cantităților astfel calculate.

Pentru inversiuni se pleacă din colțul din dreapta sus și se utilizează frecvențele din căsuțele așezate mai jos și mai la stânga față de caseta respectivă.

Se obțin astfel inversiunile:

- pentru (d): d · (e + f + g + i + j + k + m + n + o);
- pentru (c): c · (e + f + i + j + m + n);
- ...
- pentru (j): j · m.

Numărul total de inversiuni, f_i, se obține ca fiind suma cantităților astfel calculate. Restul sunt egalități care se exclud.

Și semnificația lui γ poate fi testată prin utilizarea unei formule aproximative (aplicabilă pentru $n \geq 10$):

$$z = \gamma \sqrt{\frac{f_a + f_i}{n(1 - \gamma^2)}},$$

care, în ipoteza H_0: $\gamma = 0$, este o variabilă normală redusă, putându-se aplica un test bazat pe repartiția normală standard.

Exemplul 8.3. Într-o investigație (imaginată), care a cuprins un număr de 250 de comunități umane, s-au obținut rezultatele cuprinse în tabelul următor:

consum de alcool / criminalitate	foarte înalt	înalt	scăzut	foarte scăzut
foarte înaltă	15	21	18	20
înaltă	12	9	17	16
scăzută	11	9	18	25
foarte scăzută	8	7	13	31

Pentru a decide dacă există asociere între cele două variabile considerate (consumul de alcool şi rata criminalităţii), se calculează coeficientul γ. Numărul de potriviri, după regula enunţată, este obţinut prin însumarea valorilor:

$15\times(9+17+16+9+18+25+7+13+31) = 2175$

$21\times(17+16+18+25+13+31) = 2520$

$18\times(16+25+31) = 1296$

$12\times(9+18+25+7+13+31) = 1236$

...

$18\times31 = 568.$

Se obţine $f_a = 10477$.

Pentru numărul de nepotriviri, se efectuează calculele intermediare:

$20\times(12+9+17+11+9+18+8+7+13) = 2080$

$18\times(12+9+11+9+8+7) = 1008$

...

$9\times8 = 72$

Prin însumarea acestora se obţine $f_i = 6603$.

Prin urmare,

$$\gamma = \frac{10477 - 6603}{10477 + 6603} = 0{,}227 .$$

Semnificaţia coeficientului obţinut se verifică prin calcularea statisticii:

$$z = 0{,}227 \times \sqrt{\frac{10477 + 6603}{250 \times (1 - 0{,}227^2)}} = 1{,}92 .$$

La pragul de semnificaţie $\alpha = 5\%$, se ştie că $z_{0,95} = 1{,}96$. Cum valoarea calculată este mai mică decât valoarea critică a testului, rezultă că nu putem respinge ipoteza nulă H_0: $\gamma = 0$. Concluzia este că, deşi se observă o anumită grupare preferenţială între nivelurile mari ale celor două variabile (ca şi între nivelurile inferioare), datele nu susţin ipoteza că între variabilele studiate există o asociere importantă.

Pentru a avea o imagine a sensibilităţii testului, se poate reface calculul modificând uşor datele din tabel. Dacă trei comunităţi sunt scăzute din caseta criminalitate foarte scăzută – consum de alcool foarte înalt şi adunate la prima căsuţă (criminalitate foarte înaltă – consum de alcool foarte înalt), se obţine $\gamma = 0{,}28$, $z = 2{,}40$, care arată o asociere semnificativă la pragul de semnificaţie de 5%.

8.3.3. Coeficientul de concordanţă Kendall (K)

Pornind de la două serii de date conţinând fiecare o clasificare a unei aceleiaşi serii de observaţii, coeficientul de concordanţă Kendall oferă un indicator al apropierii dintre cele două clasificări. Prin urmare, concluzia se aplică autorilor celor două clasificări exprimând gradul de similitudine existent în modul de apreciere a faptelor clasificate.

Obţinerea coeficientului Kendall necesită parcurgerea următoarelor etape:

1. se ordonează liniile tabelului de date în ordinea crescătoare a rangurilor date de primul evaluator; calculele ulterioare iau în considerare doar noua ordine a rangurilor date de cel de-al doilea evaluator;
2. se completează o coloană nouă cu numere obţinute astfel: pentru fiecare rang dat de al doilea evaluator, se numără câte ranguri mai mici decât cel considerat se află pe liniile următoare ale tabelului. Numărul obţinut se trece cu semnul minus (–), reprezentând o diferenţă de ordine faţă de clasificarea primului evaluator;
3. se completează o coloană nouă cu numere obţinute astfel: pentru fiecare rang dat de al doilea evaluator, se numără câte ranguri mai mari decât cel considerat se află pe liniile următoare ale tabelului. Numărul obţinut se trece cu semnul plus (+), reprezentând o concordanţă cu ordinea dată de primul evaluator;
4. se face suma algebrică a numerelor de pe coloanele adăugate în cei doi paşi precedenţi, obţinându-se astfel o cantitate notată cu T;
5. coeficientul lui Kendall se obţine prin următoarea formulă:

$$K = \frac{2T}{n(n-1)},$$

unde T este calculat la pasul precedent, iar n este numărul de observaţii.

Din algoritmul de calcul se observă că obţinerea coeficientului de concordanţă Kendall se bazează pe diferenţa dintre clasificări.

Coeficientul K poate lua valori între –1 (obţinut atunci când seriile de valori sunt complet inversate) şi +1 (când există concordanţă deplină între cele două clasificări), fiind astfel măsurat pe o scală direcţională (–1...+1).

Exemplul 8.4. Să presupunem că doi evaluatori sunt în situaţia de a aprecia un număr de 10 comportamente indezirabile social (comportament A – aşezarea în faţa unei cozi; comportament B – manifestare zgomotoasă la o adunare pioasă etc.). Se cere celor doi evaluatori să ordoneze comportamentele notate cu A, B, C, D, E, F, G, H, I, J, în privinţa intensităţii depreciative resimţite. Rangul 1 este atribuit comportamentului considerat cel mai departe de 0 în negativ:

Comportament	Evaluator X	Evaluator Y	–	+
a	b	c	d	e
H	1	4	–3	6
B	2	2	–1	7
G	3	1	0	7
J	4	7	–3	3
I	5	3	0	5
C	6	8	–2	2
D	7	10	–3	0
A	8	6	–1	1
E	9	5	0	1
F	10	9	0	0
Total			–13	32

Prima coloană cuprinde codificările comportamentelor indezirabile social în funcție de ordonarea primului evaluator. Rangurile atribuite de primul evaluator, ordonate ascendent, sunt trecute în coloana b. În coloana c sunt trecute rangurile atribuite de cel de-al doilea evaluator.

Coloanele d şi e, notate cu – şi +, conțin valorile calculate în paşii 2 şi 3 ai algoritmului. Explicațiile sunt pentru coloana c, singura implicată în calcule. Considerăm pe rând liniile tabelului, în ordinea naturală.

Prima linie. Comportamentul notat cu H are rangul 4 (reamintim că examinăm doar coloana c). Analizând liniile următoare din tabel, se obține:

- cu ranguri mai mici pe coloana c se găsesc comportamentele G, B şi I, notate cu 1, 2 şi 3. Existând trei ranguri mai mici, conform pasului 2 din algoritm, se trece punctajul –3 în coloana c, câte un punct negativ pentru fiecare rang mai mic;
- cu ranguri mai mari decât 4 sunt comportamentele J, C, D, A, E şi F, deci, conform pasului 3, se trece în coloana d punctajul + 6, provenit prin acordarea unui punct pentru fiecare rang mai mare decât 4.

A doua linie. Comportamentul B, cu rangul 2, are sub el un singur comportament cu rang inferior, G, deci în coloana d se trece –1. Cu ranguri mai mari există 7 comportamente, J, I, C, D, A, E şi F (se analizează doar liniile care urmează în tabel; aşadar, prima linie, H, nu mai este avută în vedere). Conform pasului 3, în coloana e se trece valoarea +7.

Se procedează în mod analog pentru liniile 3, 4, 5, ..., 10.

În final, se calculează suma algebrică a valorilor + şi a valorilor –. Se obțin astfel –13 şi 32 (a se vedea linia de total). Rezultă T = –13 + 32 = +19. Formula coeficientului de concordanță Kendall produce:

$$K = \frac{2 \times 19}{10 \times (10-1)} = \frac{38}{90} = 0{,}42.$$

Putem aprecia că între cei doi evaluatori există o concordanță bună.

8.4. Asocierea datelor continue (de interval)

În domeniul datelor continue, asocierea variabilelor este, de regulă, studiată sub denumirea de corelație (denumire care a fost introdusă şi la coeficientul de corelație a rangurilor, pentru date ordinale). Se pot distinge două mari direcții în acest studiu:

- calcularea unui indicator al asocierii (corelației);
- estimarea unui model funcțional care să reprezinte asocierea dintre variabile.

Vom schița în continuare ambele direcții, în cazul a două variabile. Generalizarea la un număr oarecare de variabile nu este conținută în această lucrare.

În 1885, Francis Galton a publicat un studiu asupra corespondenței dintre înălțimile părinților şi cele ale fiilor lor la recrutarea în armată. Datele au indicat că media înălțimilor fiilor cu părinți înalți era mai mică decât a părinților, în timp ce fiii celor mai scunzi erau, în medie, mai înalți decât părinții lor. Galton a conchis că există o tendință generală pentru regresia către media populației. Articolul său „Regresia către mediocritate în ereditate" constituie baza unei ramuri a statisticii, **analiza regresiei**, care se ocupă cu determinarea, predicția valorilor unei variabile dependente, Y, care este asociată unei variabile independente, X.

Karl Pearson a dezvoltat metoda în **analiza corelației**. În timp ce **coeficientul de regresie** reprezintă schimbarea medie a lui Y asociată cu o schimbare a lui X, **coeficientul de corelație** furnizează un indicator sintetic, direcțional, al relației de asociere dintre două variabile.

8.4.1. Dreapta de regresie

Considerăm că, într-un eşantion de volum n, s-au determinat perechile de valori (x_i, y_i), i = 1, ..., n, corespunzătoare celor două variabile pentru care dorim să studiem asocierea şi relația dintre ele.

O primă apreciere asupra distribuției comune o vom avea dacă realizăm diagrama de împrăştiere a valorilor, de fapt reprezentarea într-un sistem de axe a punctelor având coordonatele x şi y. Analiza vizuală a organizării şi formei norului de puncte obținut poate oferi indicii importante asupra relației dintre variabile.

Datele de sondaj vor susține ipoteza asocierii între variabile dacă forma norului de puncte se apropie de o curbă funcțională. Astfel, se pot aprecia asocieri liniare, curbilinii etc. Dacă în norul de puncte nu se poate distinge o tendință, se va spune că variabilele nu sunt corelate.

În figura 8.1 sunt ilustrate câteva tendințe identificabile direct.

Cazul (a) ilustrează o asociere pozitivă, (b) – o asociere negativă, (c) – lipsa de asociere, (d) – o asociere curbilinie.

Relațiile de interes pentru discuția din continuare sunt cele ilustrate în cazurile (a) şi (b), unde este identificabilă o tendință liniară în norul de puncte. Pentru a sintetiza modul în care schimbările lui Y sunt asociate cu schimbările lui X, metoda matematică utilizată este „metoda celor mai mici pătrate" (Legendre, 1806).

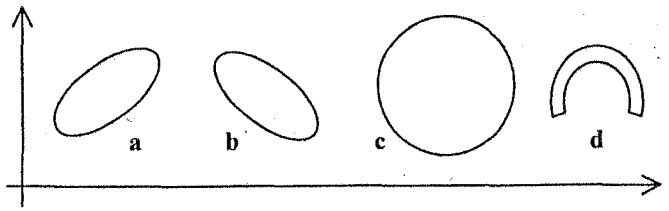


Figura 8.1. *Situații de asociere a datelor*

Aplicată în cazul nostru, asocierea dintre X şi Y este reprezentată printr-o linie dreaptă trasată printre punctele diagramei de împrăştiere, situaţie ilustrată în figura 8.2. Linia estimată este „cea mai bună" în sensul că exprimă cel mai central drum printre puncte: linia pentru care suma pătratelor distanţelor (pe verticală) dintre puncte şi dreaptă este minimă. Termenul comun pentru dreapta estimată este acela de dreapta de regresie.

Distanţele sunt desemnate ca erori (reziduuri). Deci dreapta de regresie realizează valoarea minimă a pătratelor erorilor, în sensul că orice altă dreaptă produce o sumă de pătrate mai mare. Cu alte cuvinte, putem spune că dreapta de regresie reprezintă cel mai economicos drum printre punctele observate.

Reamintim că o proprietate a mediei aritmetice este aceea că suma pătratelor diferenţelor de la medie are o valoare minimă. Astfel, se poate spune că, după cum media reprezintă punctul de echilibru pentru o distribuţie univariată de scoruri, la fel dreapta de regresie reprezintă punctul de echilibru într-o distribuţie bivariată.

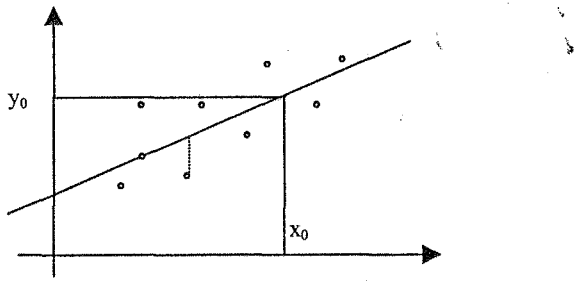


Figura 8.2. *Dreapta de regresie*

Utilitatea dreptei de regresie constă în aceea că serveşte ca bază pentru predicţia valorilor lui Y asociate valorilor lui X. Astfel, dată fiind o valoare x_0, valoarea prognozată pentru Y este y_0, ordonata pe dreapta de regresie corespunzătoare abscisei x_0 (figura 8.2). Problema estimării erorilor de predicţie va fi tratată o dată cu prezentarea modelului general.

Problema prezentată poate fi formulată matematic ca fiind determinarea cantităţilor a şi b din ecuaţia $Y_e = a + b\,X$, unde:

- Y_e este valoarea prezisă a variabilei dependente;

- a este termenul liber al dreptei de regresie (valoarea pentru X = 0);
- b este coeficientul de regresie (cantitatea cu care se modifică Y atunci când X se schimbă cu o unitate);
- X este valoarea variabilei independente.

Se demonstrează că, prin metoda celor mai mici pătrate, se obţine:

$$b = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \text{ şi}$$

$$a = \bar{y} - b \cdot \bar{x}.$$

Valoarea estimată totuşi este numai o medie la care ne putem aştepta. Acurateţea depinde de cât de bine se potriveşte dreapta de regresie cu datele reale. Această potrivire este evaluată prin considerarea unei statistici: eroarea standard a estimării, definită ca abaterea standard a erorilor de estimare (a reziduurilor estimării):

$$s = \sqrt{\frac{\sum (y_i - y_{ei})^2}{n-2}},$$

unde y_{ei} reprezintă valoarea estimată (prin ecuaţia de regresie) pentru x_i.

O eroare standard mare arată că valorile observate sunt la distanţă de dreapta de regresie şi deci aceasta este mai puţin reprezentativă pentru datele reale. În consecinţă, şi valorile prognozate sunt afectate de erori mai mari.

Valorile estimate de dreapta de regresie trebuie gândite ca medii ale valorilor Y asociate cu un X particular. Altfel spus, dacă am dispune de eşantioane repetate, mediile valorilor Y (corespunzătoare aceleiaşi valori X) vor fi date de dreapta de regresie.

Dacă notăm cu y_0 valoarea estimată corespunzătoare lui x_0, adică $y_0 = a + b\,x_0$, atunci abaterea standard a lui y_0 este

$$s(y_0) = s\sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum (x_i - \bar{x})^2}},$$

unde s este abaterea standard a estimării.

Se poate construi atunci, pentru un prag de semnificaţie fixat α, intervalul de încredere pentru valoarea estimată:

$$(y_0 - t_{1-\alpha/2;n-2}s(y_0),\ y_0 + t_{1-\alpha/2;n-2}s(y_0))$$

Atunci când valorile x_0 sunt mai depărtate de media lui X, intervalul de încredere se măreşte, iar prognoza este însoţită de erori mai mari. Predicţiile pentru valorile Y vor fi mai precise în apropierea mediei lui X.

O vizualizare a rezultatului anterior este dată în figura 8.3.

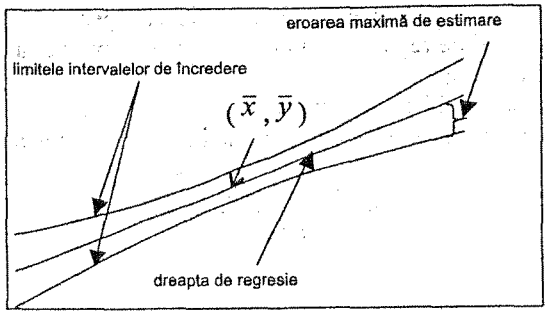
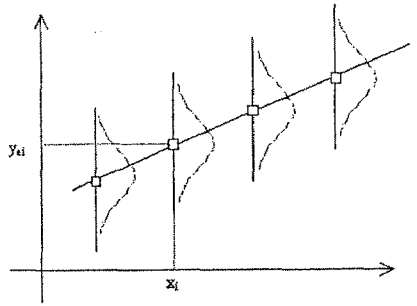


Figura 8.3. *Erorile de prognoză*

Se observă că, pentru valori depărtate de punctul central, lungimea intervalului de încredere a valorii Y prognozate se măreşte, producând o eroare de estimare mai mare.

Atunci când un model estimat va fi utilizat pentru predicţie, ar trebui satisfăcută cerinţa de **homoscedasticitate**, adică repartiţia valorilor Y să aibă (cel puţin teoretic) aceeaşi dispersie pentru toate valorile X. Pentru a putea explica mai detaliat acest aspect, se introduce noţiunea de distribuţie parţială a lui Y pentru o valoare fixată a lui X. De exemplu, are sens să discutăm despre distribuţia înălţimilor (variabila Y) copiilor cu vârsta de 12 ani (variabila X fixată la valoarea 12). Această distribuţie reprezintă distribuţia parţială a lui Y pentru X = 12. În mod analog, se poate considera distribuţia parţială a înălţimilor Y pentru vârsta X = 10 etc. Cerinţa de homoscedasticitate este satisfăcută dacă distribuţiile parţiale ale înălţimilor pentru X = 10 şi X = 12 au dispersii egale (sau aproximativ egale). Situaţia contrară se numeşte **heteroscedasticitate**. Noţiunea aceasta este valorificată ca ipoteză în alcătuirea anumitor tipuri de teste psihologice.

În figura următoare este schiţată o situaţie în care se respectă cerinţa de **homoscedasticitate**:



Se observă că, la diferite valori ale lui X, repartiţia valorilor Y este aceeaşi (în cazul considerat este o repartiţie normală), fapt sugerat prin curbele normale rotite la 90°.

8.4.2. Coeficientul de corelaţie (liniară)

Analiza de regresie constituie, în esenţă, o metodă prin care sunt permise predicţii, adică se poate estima o valoare a unei variabile Y atunci când se dispune de o valoare a variabilei asociate X. Totuşi, de multe ori în aplicaţii, cercetătorii nu sunt interesaţi sau nu cunosc care variabilă este independentă şi care dependentă în sensul cerut de practică. Ei sunt însă interesaţi să ştie dacă două variabile sunt asociate şi să cunoască gradul de asociere.

O asemenea măsură o furnizează coeficientul de corelaţie, notat r. Coeficientul de corelaţie liniară este cunoscut sub denumirea de coeficientul de corelaţie Bravais-Pearson sau, în alte lucrări, coeficientul de corelaţie Pearson. De regulă, specificarea doar a denumirii de coeficient de corelaţie este suficientă. Pentru alţi coeficienţi de corelaţie se va specifica însă numele întreg.

Coeficientul de corelaţie Bravais-Pearson are valori de la –1 la +1: o valoare nulă indică lipsa de asociere, +1 arată o asociere (corelaţie) pozitivă perfectă, o valoare de –1 arată o asociere negativă perfectă.

Coeficientul de corelaţie este definit prin:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left(\sum (x_i - \bar{x})^2\right)\left(\sum (y_i - \bar{y})^2\right)}},$$

existând, evident, şi alte formule echivalente.

Astfel, dacă se folosesc doar valorile brute, fără a se calcula mai întâi mediile şi abaterile de la medie, formula se poate scrie:

$$r = \frac{n\sum x_i y_i - (\sum x_i)(\sum y_i)}{\sqrt{[n\sum x_i^2 - (\sum x_i)^2][n\sum y_i^2 - (\sum y_i)^2]}}.$$

Pentru a aplica formulele coeficientului de corelaţie, calculele pot fi organizate după modelul tabelului următor:

Nr. crt.	x_i	y_i	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$	x_i^2	y_i^2	x_iy_i
0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	40	8	–6	–4,8	36	23,04	28,8	1600	64	320
2	45	10	–1	–2,8	1	7,84	2,8	2025	100	450
3	51	13	5	0,2	25	0,04	1,0	2601	169	663
4	38	8	–8	–3,8	64	14,44	30,04	1444	64	304
...	...	...	...	...	...	...	...	...	...	...
56	42	11	–4	–1,8	16	3,24	7,2	1764	121	462

Completarea tabelului începe cu coloanele 1 şi 2, care conţin datele brute. Se calculează mediile celor două variabile, după formula uzuală, şi apoi se pot completa restul coloanelor. Prin însumarea coloanelor 5, 6 şi 7 se obţin sumele necesare aplicării primei formule de calcul al coeficientului de corelaţie, iar prin

însumarea coloanelor 1, 2, 8, 9 şi 10 se obţin valorile necesare aplicării celei de-a doua formule.

Semnificaţia coeficientului de corelaţie poate fi testată utilizând un test Student. În aplicarea testului se presupune (este cerut) faptul că împrăştierea valorilor Y este uniformă după valorile lui X (proprietatea de homoscedasticitate). În plus, există şi ipoteze distribuţionale, de normalitate a variabilelor X şi Y. Statistica testului este

$$t = r\sqrt{\frac{n-2}{1-r^2}}.$$

În ipoteza $H_0 : \rho = 0$, statistica t este repartizată Student cu $n-2$ grade de libertate. Se respinge H_0 în favoarea ipotezei alternative nedirecţionale $H_1: \rho \neq 0$, dacă $|t| > t_{1-\alpha/2;n-2}$. Se pot considera şi teste unilaterale, după procedura generală a testelor bazate pe repartiţia Student.

Este de remarcat, după formula statisticii testului, că valoarea critică a unui coeficient de corelaţie depinde de volumul eşantionului, încât eşantioane mai mici necesită coeficienţi mai mari pentru a susţine asocierea variabilelor.

O ultimă observaţie este aceea că ipotezele testate se referă la populaţie; de aceea este vorba despre ρ (coeficientul de corelaţie în populaţie), şi nu despre r (coeficientul calculat în eşantion şi care este doar o estimaţie a coeficientului din populaţie).

***

Între cele două analize, de regresie şi de corelaţie, legătura este dată de rezultatul:

Dacă între Y şi X există o relaţie liniară perfectă, adică $Y = \alpha X + \beta$, atunci coeficientul de corelaţie este dat de

$$\rho = \frac{\alpha}{|\alpha|},$$

egalitate ce arată că $\rho = 1$, dacă $\alpha > 0$, şi $\rho = -1$, dacă $\alpha < 0$.

Reamintim din cunoştinţele de matematică faptul că funcţia dată de $Y = \alpha X + \beta$ se reprezintă grafic printr-o dreaptă orientată ascendent pentru $\alpha > 0$ şi descendent pentru $\alpha < 0$. Rezultatul prezintă o bună concordanţă cu aprecierile intuitive efectuate la începutul secţiunii despre asocieri pozitive, negative şi formele corespunzătoare ale norului de puncte observate.

Se defineşte coeficientul de determinare ca fiind pătratul coeficientului de corelaţie, r^2, şi se demonstrează că

$$r^2 = \frac{\sum_i (y_{ei} - \bar{y})^2}{\sum_i (y_i - \bar{y})^2},$$

unde prin y_{ei} se notează valorile estimate de regresia liniară, iar y_i notează valorile observate. Cum cele două sume din formulă exprimă varianţa datelor, se poate aprecia că prin coeficientul de determinare se stabileşte cât din varianţa datelor iniţiale este explicată de dreapta de regresie, adică de variabila X (cea care a fost apreciată ca independentă). Din acest motiv, coeficientul de determinare se poate exprima procentual. Astfel, un coeficient de corelaţie egal cu 0,4 produce un coeficient de determinare egal cu 0,16, având semnificaţia că doar 16% din varianţa lui Y este explicată de relaţia liniară.

Considerarea coeficientului de determinare şi a interpretării sale trebuie să mărească circumspecţia atunci când se evaluează un coeficient de corelaţie.

De asemenea, coeficienţii de corelaţie nu pot fi apreciaţi direct în privinţa proporţionalităţii. După aparenţe, o corelaţie de 0,40 este de două ori mai mică decât o corelaţie de 0,80. Comparând însă coeficienţii de determinare, adică pătratele corelaţiilor – 0,16 şi 0,64 –, se observă un raport de 1 la 4, şi nu de 1 la 2, cum apărea iniţial.

8.5. Corelaţie în psihologie – situaţii specifice

În calcularea coeficientului de corelaţie, o situaţie particulară se întâlneşte în cercetările psihologice atunci când ne aflăm în prezenţa a două şiruri de date, din care unul provine dintr-o variabilă continuă sau discretă polihotomică (asimilată unei variabile continue), iar celălalt dintr-o variabilă dihotomică. O astfel de situaţie apare la calcularea corelaţiei dintre rezultatul global al unui test şi rezultatul la un item exprimat dihotomic. O exemplificare în care se cercetează corelaţia între valorile obţinute la un item şi scorurile globale la testul care include itemul este prezentată în continuare:

Nr.crt.	X (scor global la test)	Y (scor item)
1	15	1
2	19	0
3	17	1
4	13	0
5	21	1
6	20	0
7	12	1
8	10	0
9	11	0
10	17	1
11	13	1
12	16	1
13	12	0
14	15	1
15	19	1

Corelaţia dintre aceste două serii de date se numeşte corelaţie punct biserială. Formula corelaţiei punct biseriale este următoarea:

$$r_{pbis} = \frac{\overline{X}_p - \overline{X}_q}{s_X} \times \sqrt{pq},$$

în care $\overline{X}_q$ este media aritmetică a subiecţilor care au primit la item valoarea 0, iar $\overline{X}_p$ este media aritmetică a scorurilor globale ale acelor subiecţi care au obţinut la item valoarea 1, s_X este abaterea standard a scorurilor globale, p – proporţia acelor subiecţi care au rezolvat corect itemul, $q=1-p$, deci proporţia acelor subiecţi care nu au rezolvat corect itemul. Pentru datele din tabel obţinem:

$$\overline{X}_p = \frac{15+17+21+12+17+13+16+15+19}{9} = 16{,}1$$

$$\overline{X}_q = \frac{19+13+20+10+11+12}{6} = 14{,}1,$$

p = 9/15 = 0,6, q=1–p=1–0,60=0,40, $\overline{X} = 15{,}3$, $s_X = 3{,}34$. Se obţine în final:

$$r_{pbis} = \frac{16{,}1-14{,}1}{3{,}34} \times \sqrt{0{,}60 \times 0{,}40} = 0{,}28.$$

În cazul analizei de itemi, poate cea mai rafinată analiză statistică în alcătuirea şi lansarea unui test psihologic, suntem puşi în situaţia de a stabili corelaţia dintre doi itemi. Desigur, în cazul în care corelaţia dintre doi itemi este egală cu 1 (sau foarte apropiată de 1), se poate pune întrebarea dacă ambii itemi sunt menţinuţi în cadrul unui test psihologic sau se renunţă la unul dintre ei. Raţiunile relaţionării itemilor sunt mult mai extinse şi fac obiectul de preocupare al teoriei testului psihologic. Formula pentru calcularea coeficientului φ sau a coeficientului „celor 4 câmpuri" este următoarea:

$$\varphi = \frac{p_{ij} - p_i p_j}{\sqrt{p_i q_i}\sqrt{p_j q_j}} \text{ sau } \frac{p_{ij} - p_i p_j}{\sqrt{p_i - p_i^2}\sqrt{p_j - p_j^2}}.$$

Formula se referă la doi itemi: itemul i şi itemul j. În această formulă, p_{ij} reprezintă proporţia răspunsurilor punctate cu 1 comune itemilor i şi j, p_i este proporţia răspunsurilor de 1 la itemul i, $q_i = 1 - p_i$, p_j este proporţia răspunsurilor 1 la itemul j, iar $q_j = 1 - p_j$.

Să presupunem că la un test de inteligenţă, pentru 15 subiecţi, avem următoarele rezultate la itemii i şi j:

nr.crt.	item i	item j	coincidenţă i – j
1.	1	0	
2.	1	1	+
3.	1	0	
4.	0	1	
5.	0	0	
6.	1	1	+
7.	1	0	
8.	1	0	
9.	0	1	
10.	0	1	
11.	1	0	
12.	0	0	
13.	1	1	+
14.	0	1	
15.	1	1	+

În tabel, coloanele a doua şi a treia conţin răspunsurile la itemii i şi, respectiv, j. În coloana a patra s-au completat cu semnul (+) acele observaţii în care subiecţii au răspuns de 1 la ambii itemi. Calculele necesare pentru obţinerea coeficientului ϕ sunt: p_i=9/15=0,60, p_j=8/15=0,53, p_{ij}=4/15=0,27, q_i=1–0,60=0,40, q_j=1–0,53=0,47.

Cu aceste valori se obţine

$$\varphi = \frac{p_{ij} - p_i p_j}{\sqrt{p_i q_i}\sqrt{p_j q_j}} = \frac{0{,}27 - 0{,}60 \times 0{,}53}{\sqrt{0{,}60 \times 0{,}40}\sqrt{0{,}53 \times 0{,}47}} = -0{,}20.$$

Interpretarea este similară celei date unui coeficient de corelaţie.

8.6. Corelaţii parţiale

Acest gen de corelaţie se aplică în cazul în care se studiază influenţa unei variabile atunci când aceasta acţionează în concurenţă cu alte două variabile. Problematica privind corelaţiile parţiale se înscrie într-un domeniu mai vast, al analizei multivariate. Aici se abordează doar cazul frecvent a trei variabile.

Formula corelaţiei parţiale este:

$$r_{12.3} = \frac{r_{12} - r_{13} \times r_{23}}{\sqrt{(1-r_{13}^2)(1-r_{23}^2)}},$$

unde $r_{12.3}$ este coeficientul de corelaţie parţială dintre variabilele 1 şi 2, după eliminarea influenţei variabilei 3, r_{13} şi r_{23} sunt coeficienţii de corelaţie dintre variabilele 1 şi 3 şi, respectiv, 2 şi 3.

Exemplul 8.5. În contextul în care se consideră un test 1, care diagnostichează raţionamentul spaţial, un test 2, care măsoară memoria verbală, şi vârsta (notată cu 3), coeficienţii de corelaţie fiind r(1, 2) = 0,72, r(1, 3) = 0,55), r(2, 3) = 0,65,

dorim să ştim care este valoarea corelaţiei dintre raţionamentul spaţial şi memoria verbală în condiţiile eliminării influenţei vârstei. Aplicând formula, obţinem:

$$r_{12.3} = \frac{r_{12} - r_{13} \times r_{23}}{\sqrt{(1 - r_{13}^2)(1 - r_{23}^2)}} = \frac{0,72 - 0,55 \times 0,65}{\sqrt{(1 - 0,55^2)(1 - 0,65^2)}} = 0,58,$$

care arată că între variabilele 1 şi 2, după ce s-a exclus influenţa variabilei 3, există o corelaţie mare. Exemplul 8.5 este imaginar, pur ilustrativ, în care s-au neglijat elementele de metodologie a cercetării[53].

53. Pentru o abordare completă a metodologiei cercetării psihologice, vezi Cornel Havârneanu, *Cunoaşterea psihologică a persoanei*, Editura Polirom, Iaşi, 2000.

Capitolul 9

ANALIZA DISPERSIONALĂ

Cercetările reale din psihologie, şi nu numai, implică aspecte multiple ale fenomenului studiat. Aceasta înseamnă că, pe de o parte, sunt considerate simultan mai multe caracteristici măsurate pentru obiecte, fapte, indivizi etc. Pe de altă parte, complexitatea situaţiilor reale dă posibilitatea de a considera implicate în cercetare mai multe populaţii statistice. În spiritul cercetării, sunt necesare, pe lângă descrierea fiecărei populaţii în parte, comparaţii între aceste populaţii, comparaţii care pot antrena două sau mai multe populaţii ca termeni ai operaţiei respective. Aplicaţiile statistice sunt astfel, de regulă, multivariate. Termenul **multivariat** semnifică tocmai considerarea simultană a mai multe variabile şi se distinge de termenul **univariat**, care presupune considerarea uneia singure. Într-o prelucrare statistică, dispunând de valori măsurate pentru mai multe variabile, se aplică tehnici univariate atunci când se consideră pe rând fiecare variabilă şi se prelucrează ca atare. Considerarea simultană a mai multe variabile presupune tehnici dezvoltate pentru cazul multivariat. De exemplu, asocierea a două variabile reprezintă cazul cel mai simplu de utilizare a unei tehnici statistice multivariate.

Procedurile de verificare a ipotezelor statistice implică, cel puţin în cazurile uzuale, două populaţii statistice. Acest fapt, care permite teoretic compararea oricâtor populaţii (luate două câte două), nu este în practică foarte satisfăcător. Atunci când creşte numărul de comparaţii, probabilitatea de a detecta, în mod eronat, o diferenţă semnificativă este mai mare decât nivelul de semnificaţie α fixat pentru fiecare test individual. Cu alte cuvinte, atunci când creşte numărul de comparaţii, se măreşte riscul unei erori de speţa întâi.

Să luăm, de exemplu, cazul în care se studiază un grup de subiecţi, care sunt caracterizaţi de două variabile dihotomice: bărbat/femeie şi urban/rural. Dintr-o dată, apar patru subgrupuri ($= 2 \times 2$): bărbat + urban, bărbat + rural, femeie + urban, femeie + rural. Comparaţiile între aceste grupuri sunt în număr de 6 ($= C_4^2$). Dacă la o comparaţie se alege ca nivel de semnificaţie $\alpha = 0,05$ şi se presupune că testele sunt independente între ele, atunci probabilitatea de a nu respinge ipoteza nulă în nici un test dintre cele 6, dacă H_0 este adevărată, este:

$$P(\text{nu se respinge } H_0 \text{ în nici un test} \mid H_0 \text{ este adevărată}) = (0,95)^6 = 0,7351.$$

Prin urmare, probabilitatea evenimentului complementar, de a respinge H_0, deşi aceasta este adevărată, este:

$$P(\text{se respinge } H_0 \text{ în cel puţin un test} \mid H_0 \text{ este adevărată}) = 1 - (0{,}95)^6 = 1 - 0{,}7351 = 0{,}2649.$$

Se poate spune că există o şansă de 26% de respingere a ipotezei nule, deşi ea este adevărată. Reamintim că, prin alegerea lui $\alpha = 0{,}05$, şansa era de 5%. Se observă astfel că în cazul efectuării mai multor teste de comparare creşte probabilitatea de a accepta o diferenţă semnificativă acolo unde ea nu există.

Apare astfel necesitatea de a dezvolta proceduri noi, care să nu mărească riscul apariţiei unor erori de speţa întâi, pentru analiza situaţiilor care presupun comparaţii multiple. **Analiza dispersională** este denumirea generică pentru o clasă întreagă de astfel de tehnici, dezvoltate pentru situaţiile în care influenţele unor variabile (factori) se însumează, influenţa globală putând fi evidenţiată ca parte a variaţiei unei variabile dependente.

Alt termen pentru analiza dispersională este **analiza varianţei**. Este frecventă utilizarea prescurtării ANOVA (din ANalysis Of VAriance).

Termeni specifici. Existenţa mai multor populaţii (grupuri) într-o cercetare este determinată de una sau mai multe variabile care fac distincţia între populaţii: sex, mediu de reşedinţă, nivel de pregătire, nivel de participare etc. O astfel de variabilă este numită **variabilă independentă**. Variabila după care se compară grupurile este numită **variabilă dependentă**. Într-o experimentare psihologică, variabila independentă este cea manipulată de experimentator, iar cea dependentă este variabila observată de către acesta.

O variabilă independentă mai este numită **factor**, pentru că se consideră a fi unul dintre factorii care influenţează comportarea variabilei dependente. Trebuie totuşi, în cazul utilizării acestei denumiri, să se facă diferenţierea de analiza factorială, ce are în statistică o semnificaţie proprie. În analiza factorială, factorul este o variabilă ipotetică susceptibilă de a da explicaţii pentru corelaţiile dintre variabilele observate (factor comun).

Unii cercetători utilizează pentru o variabilă independentă şi denumirea de **efect principal**, deoarece se urmăreşte efectul variabilei independente asupra variabilei dependente. Denumirea nu este foarte fericită, deoarece duce la ideea de cauzalitate, care nu poate fi detectată numai prin metode statistice. Discuţia asupra relaţiei cauză – efect este mult mai vastă şi este schiţată în capitolul dedicat asocierii dintre variabile.

După numărul variabilelor independente, al factorilor, se distinge analiza dispersională unifactorială, bifactorială etc., fiecare prezentând probleme specifice. În continuare sunt ilustrate cazurile cu unul sau doi factori, atât pentru că sunt mai simple ca aparat statistic, cât şi pentru motivul că analizele (de acest tip) cu mai mult de doi factori, posibile teoretic, sunt foarte dificil de interpretat în termeni reali.

Exemplele constituie subiectul unei secţiuni separate.

9.1. Analiza dispersională unifactorială (ANOVA pe o cale)

9.1.1. Structura datelor

Valorile unei caracteristici X sunt măsurate în k eşantioane independente, obţinute din k populaţii. Populaţiile se iau repartizate normal, cu mediile $\mu_1, \mu_2, \ldots, \mu_k$ şi dispersiile egale $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2$. Se pune problema dacă populaţiile pot fi considerate identice din punctul de vedere al caracteristicii X.

Pentru verificarea ipotezei de egalitate a dispersiilor, se va utiliza unul dintre testele prezentate în secţiunea finală a acestui capitol. În cazul în care ipoteza normalităţii nu este îndeplinită (sau nu este verificată), erorile care se introduc nu sunt importante atunci când volumul eşantioanelor este suficient de mare (> 4), iar numărul grupurilor este relativ mic.

Formal, dispunem de k grupuri de valori, de volume $n_1, n_2, \ldots, n_k$, notate cu $A_1, A_2, \ldots, A_k$. În sensul dezbaterii generale, se poate considera că $A_1, A_2, \ldots$ sunt categoriile unei variabile (de obicei nominale) care fac deosebirea dintre grupuri. Aceasta este variabila independentă a analizei. Variabila X este variabila dependentă, după care se compară populaţiile. Schematic, situaţia este prezentată în tabelul 9.1:

Tabelul 9.1. *Organizarea datelor de sondaj pentru analiza dispersională unifactorială*

A_1	...	A_i	...	A_k
x_{11}	...	x_{i1}	...	x_{k1}
x_{12}	...	x_{i2}	...	x_{k2}
...	...		...	
x_{1n_1}	...	x_{in_i}	...	x_{kn_k}

Reamintim că baza oricărei aplicaţii statistice o constituie variaţia valorilor unei caracteristici (pentru un şir constant de valori, nu prezintă interes nici o prelucrare). În cazul structurării datelor într-un tabel similar tabelului 9.1, se pune întrebarea cât din variaţia datelor poate fi atribuită apartenenţei la grupuri diferite (de exemplu: *este de aşteptat ca x_{12} să fie diferit de x_{34} întrucât aparţin unor grupuri diferite*). Aceasta este, în sensul discuţiei din primul capitol, variaţia sistematică a datelor, împărţirea în grupuri provenind de obicei din aceea că statisticianul se aşteaptă să apară diferenţe între grupuri (în cazul contrar, dacă statisticianul nu s-ar aştepta la

asemenea diferenţe, nu ar avea sens să le marcheze la strângerea datelor). Este evident atunci că o asemenea variaţie în date poate fi evidenţiată doar dacă nu este „acoperită" de sursele de variaţie neluate în considerare (cu alte cuvinte, dacă sursa de variaţie a împărţirii în grupuri este suficient de puternică). Prin metoda analizei dispersionale:

- se evidenţiază variaţia datelor datorată împărţirii în grupuri;
- se compară cu variaţia totală a datelor;
- dacă se poate decide că sursa de variaţie a împărţirii în grupuri este suficient de puternică, atunci se consideră că populaţiile nu sunt omogene; dacă sursa de variaţie respectivă este nesemnificativă, atunci se consideră că populaţiile sunt omogene.

Variaţia datorată diferenţelor dintre grupuri este definită ca **variaţie explicată** – partea din variaţia variabilei dependente explicată de variabila independentă. Cu alte cuvinte, variaţia explicată este partea explicată de către împărţirea în grupuri. Variaţia, care rămâne după separarea variaţiei explicate, se defineşte ca fiind **variaţie reziduală (variaţia neexplicată)** şi se datorează unor surse întâmplătoare de variaţie. Variaţia explicată mai este denumită şi **variaţie între grupuri** (variaţia intergrupuri, exterioară), iar cea reziduală – **variaţie în grupuri** (variaţia intragrupuri, internă).

Rezumând, se poate considera că modelul de bază al analizei dispersionale afirmă că orice valoare a variabilei X este obţinută prin cumularea a două efecte, unul sistematic şi unul întâmplător:

x_{ij} = efect sistematic (al nivelului A_i) + efect întâmplător (din nivelul A_i).

Deoarece expresia matematică este mai complicată, depinzând şi de alte presupuneri asupra variabilei independente, explicitarea expresiei se face în altă secţiune a capitolului.

Pe baza datelor din tabelul 9.1 se definesc mediile grupurilor (pentru fiecare se consideră doar valorile grupului):

$$\bar{x}_i = \frac{x_{i1} + x_{i2} + \ldots + x_{in_i}}{n_i} = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}\ ,\ i = 1,2,\ldots,k$$

şi media generală (luând în considerare grupurile reunite):

$$\bar{x} = \frac{1}{n_1 + n_2 + \ldots + n_k}\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij}\ .$$

Se mai notează numărul total de observaţii:

$$n = n_1 + n_2 + \ldots + n_k\ .$$

Cu aceste notaţii, se defineşte variaţia din interiorul unui grup prin abaterile faţă de media grupului, iar variaţia dintre grupuri prin abaterile de la media generală a mediilor grupurilor. Se demonstrează că are loc relaţia:

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - \bar{x})^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^2 + \sum_{i=1}^{k}\sum_{j=1}^{n_i}(\bar{x}_i - \bar{x})^2\ .$$

Suma din partea stângă reflectă variaţia globală a variabilei X (este suma pătratelor abaterilor, întâlnită şi la calculul dispersiei). Prima sumă din partea dreaptă a egalităţii reflectă variaţia valorilor în interiorul grupurilor (abaterea unei valori se calculează faţă de media grupului căruia îi aparţine valoarea), iar ultima sumă reflectă variaţia dintre grupuri (se bazează pe abaterile de la media generală a mediilor grupurilor). Notăm:

$$SP_g = \sum_i\sum_j (x_{ij} - \bar{x})^2$$

$$SP_{exp} = \sum_i\sum_j (\bar{x}_i - \bar{x})^2 = \sum_i n_i(\bar{x}_i - \bar{x})^2$$

$$SP_{rez} = \sum_i\sum_j (x_{ij} - \bar{x}_i)^2$$

unde denumirile uzuale ale analizei dispersionale sunt: SP_g este suma pătratelor globală, SP_{exp} este suma pătratelor explicată, iar SP_{rez} este suma pătratelor reziduală. Prin urmare:

$$SP_g = SP_{exp} + SP_{rez},$$

cu interpretarea de descompunere a variaţiei globale în variaţia explicată şi variaţia reziduală.

Cele trei sume de pătrate au următoarele grade de libertate, notate corespunzător fiecărei sume:

$$\nu_g = \sum_{i=1}^{k} n_i - 1 = n - 1,\ \ \nu_{exp} = k - 1,\ \ \nu_{rez} = n - k.$$

Prin raportarea unei sume de pătrate la numărul ei de grade de libertate, se obţine media pătratică. Astfel:

$$s_g^2 = \frac{SP_g}{\nu_g};\ \ s_{exp}^2 = \frac{SP_{exp}}{\nu_{esp}};\ \ s_{rez}^2 = \frac{SP_{rez}}{\nu_{rez}},$$

care, potrivit discuţiei de la introducerea dispersiei ca indicator al împrăştierii, caracterizează împrăştierile globală, explicată şi, respectiv, reziduală.

Dintre acestea, s_{rez}^2 estimează absolut corect dispersia teoretică σ^2 (valoarea comună a dispersiilor populaţiilor considerate), indiferent dacă mediile de sondaj ale grupurilor estimează sau nu aceeaşi medie teoretică μ.

Dacă mediile de sondaj sunt omogene, atunci şi celelalte medii pătratice, s_g^2 şi s_{exp}^2, sunt estimaţii absolut corecte ale dispersiei comune σ^2. Dacă mediile de sondaj nu sunt omogene (ca urmare a influenţei factorului sistematic de clasare în grupuri), s_g^2 şi s_{exp}^2 nu vor mai estima absolut corect dispersia comună. Aceste rezultate arată că între s_{rez}^2 şi s_{exp}^2 apar diferenţe semnificative atunci când mediile grupurilor sunt neomogene (influenţate de efectele variabilei independente), deci compararea acestor cantităţi printr-un test F evidenţiază efectul variabilei independente. Din acest motiv, analiza dispersională este uneori desemnată printr-un test F pentru o comparare multiplă.

Cu datele calculate potrivit formulelor prezentate se completează tabelul 9.2 al analizei dispersionale unifactoriale, denumit frecvent tabel ANOVA. Deşi sunt interpretabile toate informaţiile din acest tabel, scopul principal este obţinerea valorii calculate F.

Tabelul 9.2. *Tabelul analizei dispersionale unifactoriale (ANOVA)*

Sursa variaţiei	Suma pătratelor	Grade de libertate	Media pătratelor	F
Între grupuri (externă)	SP_{exp}	ν_{exp}	s_{exp}^2	$F = \frac{s_{exp}^2}{s_{rez}^2}$
În grupuri (internă)	SP_{rez}	ν_{rez}	s_{rez}^2	
Globală	SP_g	ν_g	s_g^2	

Testul F. Compararea cantităţilor s_{exp}^2 şi s_{rez}^2 se efectuează într-un test F cu ipotezele:

$$\begin{cases} H_0 : \mu_1 = \mu_2 = \ldots = \mu_k \text{ (mediile grupurilor sunt egale);} \\ H_1 : \text{există } \mu_t \neq \mu_j \text{ (cel puţin două medii sunt inegale).} \end{cases}$$

De remarcat că în ipoteza alternativă se afirmă existenţa unei diferenţe între două medii, dar nu se pot specifica mediile diferite; nu se afirmă că toate mediile sunt diferite două câte două.

În condiţiile ipotezei H_0, cantitatea F calculată în tabelul ANOVA urmează o repartiţie Fisher-Snedecor cu ν_{exp} şi ν_{rez} grade de libertate, $F_{\nu_{exp};\nu_{rez}}$, ceea ce permite efectuarea unui test statistic.

Decizia. Dacă, pentru un prag de semnificaţie α fixat, valoarea F calculată este mai mare sau egală cu quantila corespunzătoare a repartiţiei F,

$$F \geq F_{1-\alpha;\nu_{exp};\nu_{rez}},$$

atunci se respinge ipoteza nulă în favoarea ipotezei alternative. Cu alte cuvinte, în acest caz se poate accepta ipoteza că mediile grupurilor nu sunt omogene; există cel puţin două medii diferite între ele. Se conchide că diferenţele dintre grupuri justifică o parte semnificativă a variaţiei variabilei dependente.

În caz contrar, nu se respinge ipoteza nulă a omogenităţii mediilor: împărţirea în grupuri nu este semnificativă (grupurile sunt similare în privinţa variabilei dependente).

> **Observaţie.** Reamintim presupunerea iniţială că toate grupurile sunt caracterizate de aceeaşi dispersie.

9.1.2. Modele ale analizei dispersionale unifactoriale

În funcţie de modul de stabilire a grupurilor, pot să apară două situaţii, după cum sunt luate în considerare sau nu toate grupurile posibile de interes:

1. Grupurile corespund categoriilor unei scale nominale (ordinale), toate categoriile scalei fiind luate în considerare. În acest caz, se spune că se studiază întreaga populaţie a surselor. Situaţia apare şi atunci când, prin convenţie, se *consideră* că se studiază întreaga populaţie a surselor. În acest caz, se compară grupurile între ele şi se obţin concluzii valabile pentru fiecare grup în parte. Acesta este **modelul cu efecte sistematice** (sau **modelul 1**).
2. Grupurile corespund unor surse alese întâmplător dintr-o populaţie a surselor. În acest caz, este important ca, pe lângă compararea grupurilor, să se obţină informaţii despre mulţimea tuturor surselor. Este ca şi cum, din multitudinea de categorii ale unei scale nominale (ordinale), se aleg la întâmplare câteva categorii, se obţin eşantioane din grupurile corespunzătoare categoriilor selectate, iar prelucrarea se efectuează asupra acestor eşantioane. Acesta este **modelul cu efecte întâmplătoare** (sau **modelul 2**).

De exemplu, într-o situaţie imaginată, să presupunem că într-un birou sunt 5 operatoare şi sunt supuse toate unei investigări, cu înregistrarea valorilor la momente aleatorii în timp, pe parcursul unei săptămâni. Pentru a stabili omogenitatea celor cinci grupe de valori observate, cum s-au luat în considerare *toate* operatoarele, se va utiliza modelul cu efecte sistematice.

Dacă într-o hală industrială sunt 100 de operatoare, dintre acestea se aleg 8 la întâmplare şi doar acestea sunt supuse investigării; atunci modelul adecvat este modelul cu efecte întâmplătoare, întrucât cele 8 grupe de valori s-au obţinut în mod aleatoriu dintr-o sută posibile, iar rezultatele trebuie să se refere la toate cele 100 de operatoare, şi nu numai la cele 8 din grupul experimental.

Denumirile modelelor sunt explicate de utilizarea termenului *efect* pentru categoriile variabilei independente. Este de remarcat faptul că, uneori, pe acelaşi set de date se aplică ambele modele, în funcţie de scopurile urmărite în cercetare.

> **Observaţie.** Indiferent de modelul aplicat, calculele care conduc la stabilirea tabelului ANOVA sunt aceleaşi. Diferenţele dintre modele apar la concluzii şi la obţinerea altor estimaţii.

9.1.2.1. Modelul cu efecte sistematice

În acest model, fiecare grup (populație) este considerat separat, ca o entitate de interes în sine pentru cercetător, şi nu ca un element afectat de erori de sondaj, element extras aleatoriu dintr-o mulțime posibilă. Fiecare valoare de sondaj, x_{ij}, este obținută atunci ca fiind suma dintre valoarea medie a grupului (populației) căruia îi aparține, μ_i, şi o abatere (eroare) normal repartizată, ε_{ij}, ceea ce este formalizat prin:

$$\begin{cases} x_{ij} = \mu_i + \varepsilon_{ij}, \quad 1 \le i \le k,\ 1 \le j \le n_i; \\ \varepsilon_{ij} \sim N(0,\sigma^2), \text{ v.a. independente.} \end{cases}$$

Este de remarcat faptul că erorile sunt proprii fiecărei observații, sunt independente între ele, dar toate sunt repartizate cu aceeaşi dispersie σ^2. Aceste erori includ atât erorile sistematice, cât şi erorile aleatorii datorate surselor de variație care nu au fost avute în vedere (sau sunt necunoscute). Considerarea unei aceleiaşi dispersii exprimă matematic situația experimentală în care grupurile sunt subpopulații ale unei populații mai vaste.

Verificarea ipotezei nule $H_0 : \mu_1 = \mu_2 = \ldots = \mu_k = \mu$ contra alternativei H_1 : *există o pereche* $\mu_i \neq \mu_j$ permite concluzii asupra omogenității grupurilor.

În cazul în care H_0 nu se respinge, $F < F_{1-\alpha;\nu_{exp};\nu_{rez}}$, atunci cele k grupuri pot fi considerate extrase din aceeaşi populație normală cu media μ şi dispersia σ^2, dispersie estimată de

$$s_g^2 = \frac{SP_g}{\nu_g} = \frac{SP_{exp} + SP_{rez}}{\nu_{exp} + \nu_{rez}},$$

astfel încât intervalele de încredere pentru μ şi σ^2 pot fi construite după formulele uzuale:

$$\bar{x} - t_{1-\alpha/2;n-1}\frac{s_g}{\sqrt{n}} \le \mu \le \bar{x} + t_{1-\alpha/2;n-1}\frac{s_g}{\sqrt{n}}$$

$$\frac{n-1}{\chi^2_{\nu_g;1-\alpha/2}}s_g^2 \le \sigma^2 \le \frac{n-1}{\chi^2_{\nu_g;\alpha/2}}s_g^2.$$

Dacă ipoteza nulă se respinge, atunci mediile grupurilor nu pot fi considerate omogene şi pentru calcularea intervalelor de încredere se va utiliza estimația s^2_{rez} a lui σ^2:

$$\bar{x} - t_{1-\alpha/2;n-k}\frac{s_{rez}}{\sqrt{n}} \le \mu \le \bar{x} + t_{1-\alpha/2;n-k}\frac{s_{rez}}{\sqrt{n}},$$

$$\bar{x_i} - t_{1-\alpha/2;n-k}\frac{s_{rez}}{\sqrt{n_i}} \le \mu_i \le \bar{x_i} + t_{1-\alpha/2;n-k}\frac{s_{rez}}{\sqrt{n_i}},$$

$$\frac{n-k}{\chi^2_{n-k;1-\alpha/2}}s^2_{rez} \le \sigma^2 \le \frac{n-k}{\chi^2_{n-k;\alpha/2}}s^2_{rez},$$

unde $i = 1, 2, \ldots, k$.

Pentru diferența dintre media unui grup şi media populației globale (în cazul în care grupurile sunt considerate subpopulații ale unei populații de bază), se utilizează formula:

$$(\bar{x_i} - \bar{x}) - t_{1-\alpha/2;n-k}\frac{s_{rez}}{\sqrt{\frac{n_i \cdot n}{n+n_i}}} \le \mu_i - \mu \le (\bar{x_i} - \bar{x}) + t_{1-\alpha/2;n-k}\frac{s_{rez}}{\sqrt{\frac{n_i \cdot n}{n+n_i}}}.$$

9.1.2.2. Modelul cu efecte întâmplătoare

Cum se presupune că grupurile, în cadrul acestui model, ar fi extrase aleatoriu dintr-o populație de grupuri, are sens să considerăm că fiecare observație este obținută ca fiind suma dintre media generală a populației de bază, efectul nivelului corespunzător al variabilei independente şi abaterea (eroarea) de la media grupului. Cu alte cuvinte, față de modelul cu efecte sistematice, media grupului nu mai este o constantă, ci o variabilă aleatorie măsurată prin abaterea de la media populației. Ambele abateri care s-au introdus sunt considerate normal repartizate.

Relațiile matematice corespunzătoare sunt:

$$\begin{cases} x_{ij} = \mu + \delta_i + \varepsilon_{ij}, \quad i = \overline{1,k},\ j = \overline{1,n_i} \\ \delta_i \sim N(0,\sigma^2_\delta) \text{ sunt v.a. independente, } i = \overline{1,k} \\ \varepsilon_{ij} \sim N(0,\sigma^2) \text{ sunt v.a. independente, } i = \overline{1,k},\ j = \overline{1,n_i} \end{cases}$$

Se observă că $\mu + \delta_i$ reprezintă echivalentul mediei grupului din modelul cu efecte sistematice. Egalitatea mediilor grupurilor presupune atunci verificarea ipotezei $H_0 : \delta_1 = \delta_2 = \ldots = \delta_k = 0$, contra ipotezei alternative că nu toate efectele δ sunt egale cu zero.

Testul statistic se bazează tot pe tabelul ANOVA şi, în cazul în care ipoteza nulă nu este respinsă, $F < F_{1-\alpha;\nu_{exp};\nu_{rez}}$, continuarea este aceeaşi ca la modelul cu efecte sistematice.

În cazul în care se respinge ipoteza nulă, se poate obține o estimație a dispersiei abaterilor δ_i:

$$s_\delta^2 = \frac{s_{ext}^2 - s_{rez}^2}{n^*}, \text{ unde } n^* = \frac{1}{k-1}\left(n - \frac{1}{n}\sum_{i=1}^{k} n_i^2\right),$$

ceea ce permite judecăți probabiliste asupra abaterilor grupurilor de la media generală. Reamintim că prin n se notează suma efectivelor grupurilor (volumul total). Dacă grupurile sunt echilibrate, adică au volume egale, $n_1 = n_2 = \ldots = n_k$, atunci n* este egal cu volumul comun.

Pentru intervalul de încredere al mediei globale se poate utiliza formula aproximativă (valabilă atunci când populația efectelor este finită, de volum K):

$$\bar{x} - t_{1-\alpha/2;\,k-1}\sqrt{\frac{K-k}{K-1}\cdot\frac{s_\delta^2}{k} + \frac{s_{rez}^2}{n}} \le \mu \le \bar{x} + t_{1-\alpha/2;\,k-1}\sqrt{\frac{K-k}{K-1}\cdot\frac{s_\delta^2}{k} + \frac{s_{rez}^2}{n}},$$

populația surselor considerându-se finită dacă se selectează mai mult de o zecime dintre efectele posibile.

Pentru o populație de efecte infinită (foarte numeroasă), intervalul de încredere al mediei globale este dat de:

$$\bar{x} - t_{1-\alpha/2;\,k-1}\sqrt{\frac{s_\delta^2}{k} + \frac{s_{rez}^2}{n}} \le \mu \le \bar{x} + t_{1-\alpha/2;\,k-1}\sqrt{\frac{s_\delta^2}{k} + \frac{s_{rez}^2}{n}}.$$

9.2. *Analiza dispersională bifactorială (ANOVA pe două căi)*

Presupunem că suntem interesați să comparăm între ele grupuri (populații statistice) caracterizate de două atribute. De exemplu, să comparăm atenția unor grupuri caracterizate de nivelul de oboseală și consumul de alcool. Scopul propus al unei asemenea cercetări poate fi acela de a stabili:

- dacă atenția depinde de oboseală;
- dacă atenția depinde de consumul de alcool;
- dacă atenția celor odihniți depinde de consumul de alcool într-un mod diferit de al celor obosiți (efectul conjugat al celor doi factori) etc.

Primele două categorii de obiective pot fi determinate printr-o analiză dispersională unifactorială. Al treilea tip de obiectiv implică însă considerarea simultană a celor doi factori de diferențiere a grupurilor, ceea ce nu mai permite utilizarea analizei unifactoriale.

Din această abordare sumară se poate observa că implicarea a doi factori necesită un instrument de analiză care să poată diferenția influența primului factor, influența celui de-al doilea factor, precum și influența combinată, interacțiunea celor doi factori. Din acest motiv, analiza dispersională bifactorială nu este o simplă juxtapunere a două analize unifactoriale.

Metoda se bazează tot pe descompunerea variației totale în variație explicată și variație reziduală, situație care este ilustrată în figura 9.1. Se observă modurile de explicare a variației: de fiecare factor (individual) sau de interacțiunea celor doi factori. După descompunerea variației totale, decizia se ia pe baza comparării diferitelor „părți" ale variației explicate cu variația reziduală. Dacă o parte explicată este semnificativ mai mare decât partea reziduală, atunci se va considera că respectivul factor sau interacțiunea lor prezintă importanță pentru nivelul valorii variabilei dependente.

Datorită prezenței interacțiunilor, creșterea numărului de factori complică foarte mult interpretarea rezultatelor. Pentru trei factori, A, B, C, apar interacțiunile AB, AC, BC, dar și interacțiunea ABC. Se observă astfel că numărul interacțiunilor, 4, este mai mare decât numărul inițial de factori. Din acest motiv, utilizarea analizei dispersionale pentru mai mult de doi factori este suficient de redusă. Trebuie notat faptul că introducerea unui nou factor necesită formule de calcul specifice.

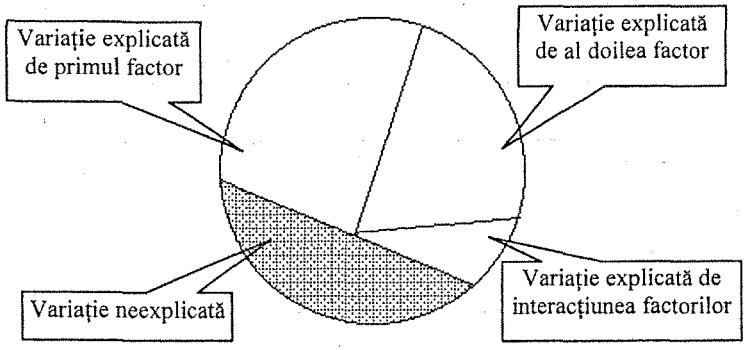

Variație explicată de primul factor
Variație explicată de al doilea factor
Variație neexplicată
Variație explicată de interacțiunea factorilor

Figura 9.1. *Descompunerea variației globale în analiza dispersională bifactorială*

9.2.1. Structura datelor

Notăm cu A și B variabilele independente (factorii), iar cu X variabila dependentă (după care se compară grupurile determinate de nivelurile factorilor). Presupunem că factorul A are h niveluri, iar factorul B are q niveluri. Rezultă că factorii A și B determină hq grupuri. Se obișnuiește ca valorile de sondaj corespunzătoare grupurilor să fie aranjate într-o structură dreptunghiulară, cu hq celule, fiecare eșantion ocupând celula corespunzătoare nivelurilor A_i și B_j, care identifică grupul. O asemenea structură este desemnată prin noțiunea de **plan bifactorial**. Deși este posibil ca eșantioanele să aibă volume diferite, vom considera doar cazul în care toate celulele conțin un număr egal de valori, notat cu n, caz în care se spune că avem un **plan factorial echilibrat**.

Tabelul 9.3 conţine structura generală de date pentru o analiză dispersională cu doi factori. Valorile variabilei dependente X sunt indexate cu trei indici: primul pentru nivelul factorului A, al doilea pentru nivelul factorului B, iar ultimul pentru numerotarea în cadrul celulei. Deci prin x_{ijk} se notează a k-a valoare din eşantionul care corespunde grupului determinat de nivelul i al factorului A şi nivelul j al factorului B.

Tabelul 9.3. *Structura de date pentru analiza dispersională bifactorială*

A \ B	B_1	...	B_j	...	B_q
A_1	x_{111}	...	x_{1j1}	...	x_{1q1}
	⋮	⋮	⋮	⋮	⋮
	x_{11n}	...	x_{1jn}	...	x_{1qn}
⋮	⋮	⋮	⋮	⋮	⋮
A_i	x_{i11}	...	x_{ij1}	...	x_{iq1}
	⋮	⋮	⋮	⋮	⋮
	x_{i1n}	...	x_{ijn}	...	x_{iqn}
⋮	⋮	⋮	⋮	⋮	⋮
A_h	x_{h11}	...	x_{hj1}	...	x_{hq1}
	⋮	⋮	⋮	⋮	⋮
	x_{h1n}	...	x_{hjn}	...	x_{hqn}

Pentru această structură de date trebuie remarcată organizarea în linii şi coloane a *celulelor*. Se va putea vorbi astfel, la un nivel mai general, de linia nivelului A_i sau de coloana nivelului B_j.

Introducem următoarele notaţii pentru diferitele tipuri de valori medii care se pot defini:

- **media globală** (media tuturor valorilor observate pentru variabila X, indiferent de nivelurile factorilor):

$$\bar{x} = \frac{1}{hqn}\sum_{i=1}^{h}\sum_{j=1}^{q}\sum_{k=1}^{n} x_{ijk};$$

- **media celulei** (i, j) (media eşantionului determinat de nivelurile A_i şi B_j), definită pentru orice i şi orice j:

$$\bar{x}_{ij} = \frac{1}{n}\sum_{k=1}^{n} x_{ijk};$$

- **media liniei** i (media valorilor observate la nivelul i al factorului A, indiferent de nivelul factorului B), definită pentru orice i:

$$\bar{x}_{i*} = \frac{1}{qn}\sum_{j=1}^{q}\sum_{k=1}^{n} x_{ijk} = \frac{1}{q}\sum_{j=1}^{q}\bar{x}_{ij};$$

- **media coloanei** j (media valorilor observate la nivelul j al factorului B, indiferent de nivelul factorului A), definită pentru orice j:

$$\bar{x}_{*j} = \frac{1}{hn}\sum_{i=1}^{h}\sum_{k=1}^{n} x_{ijk} = \frac{1}{h}\sum_{i=1}^{h}\bar{x}_{ij}.$$

Pentru media globală se mai utilizează şi notaţia $\bar{x}_{**}$, similară notaţiilor utilizate pentru mediile liniilor şi coloanelor. Au loc şi relaţiile:

$$\bar{x}\ (\text{sau } \bar{x}_{**}) = \frac{1}{hq}\sum_{i=1}^{h}\sum_{j=1}^{q}\bar{x}_{ij} = \frac{1}{h}\sum_{i=1}^{h}\bar{x}_{i*} = \frac{1}{q}\sum_{j=1}^{q}\bar{x}_{*j}.$$

Fiecare dintre mediile calculate este interpretabilă în sine, ca orice valoare medie calculată pentru un eşantion. Se pot utiliza, atunci când are sens, teste individuale de comparare a acestor valori medii, două câte două (fără a implica în vreun fel variabilele independente A şi B).

Utilizând respectivele valori medii, se demonstrează că are loc relaţia următoare de descompunere a variaţiei totale în variaţia datorată factorului A, variaţia datorată factorului B, variaţia datorată interacţiunii AB şi variaţia reziduală:

$$\begin{aligned}
SP_g &= \sum_i\sum_j\sum_k (x_{ijk} - \bar{x})^2 = \\
&= \sum_i\sum_j\sum_k (x_{i*} - \bar{x})^2 + \sum_i\sum_j\sum_k (x_{*j} - \bar{x})^2 + \\
&+ \sum_i\sum_j\sum_k (\bar{x}_{ij} - \bar{x}_{i*} - \bar{x}_{*j} + \bar{x})^2 + \sum_i\sum_j\sum_k (x_{ijk} - \bar{x}_{ij})^2 = \\
&= SP_A + SP_B + SP_{AB} + SP_{rez},
\end{aligned}$$

unde însumările după i, j şi k se fac până la limitele menţionate deja, h, q şi n, iar notaţiile SP_A, SP_B şi SP_{AB} ale sumelor de pătrate sunt în ordinea apariţiei lor. Dacă

se ţine cont de faptul că, în sumele respective, cantitatea însumată este constantă în raport cu unii indici de însumare, se obţin formulele simplificate:

$$SP_A = qn\sum_{i=1}^{h}(\bar{x}_{i*} - \bar{x})^2$$

$$SP_B = hn\sum_{j=1}^{q}(\bar{x}_{*j} - \bar{x})^2$$

$$SP_{AB} = n\sum_{i=1}^{h}\sum_{j=1}^{q}(\bar{x}_{ij} - \bar{x}_{i*} - \bar{x}_{*j} + \bar{x})^2$$

$$SP_{rez} = \sum_{i=1}^{h}\sum_{j=1}^{q}\sum_{k=1}^{n}(x_{ijk} - \bar{x}_{ij})^2,$$

unde:

- SP_A este suma pătratelor datorate factorului A şi măsoară variaţia explicată de acest factor. Suportul intuitiv este imediat, prin aceea că se însumează abaterile pătrate ale mediilor liniilor de la media generală;
- SP_B este suma pătratelor datorate factorului B şi măsoară variaţia explicată de acest factor. Suportul intuitiv este asigurat de însumarea abaterilor pătrate ale mediilor coloanelor de la media generală;
- SP_{AB} este suma pătratelor datorate interacţiunii factorilor A şi B; măsoară variaţia explicată de această interacţiune. Suportul intuitiv este evident dacă observăm că:

$$\bar{x}_{ij} - \bar{x}_{i*} - \bar{x}_{*j} + \bar{x} = (\bar{x}_{ij} - \bar{x}) - (\bar{x}_{i*} - \bar{x}) - (\bar{x}_{*j} - \bar{x}),$$

  adică din abaterea mediei celulei faţă de media generală se scade efectul liniei şi efectul coloanei, ceea ce rămâne fiind definit ca efectul interacţiunii factorilor;
- SP_{rez} este suma reziduală a pătratelor şi măsoară variaţia neexplicată de cei doi factori sau de interacţiunea lor. Intuitiv, abaterile observaţiilor de la media grupului arată variaţia internă, proprie grupului.

Se poate spune că $S_A + S_B + S_{AB}$ măsoară variaţia explicată de factorii A şi B – variaţia exterioară, între grupuri –, în timp ce S_{rez} măsoară variaţia neexplicată, variaţia internă a grupurilor.

Sumele de pătrate au gradele de libertate:

$$\nu_g = hqn - 1, \quad \nu_A = h - 1, \quad \nu_B = q - 1,$$
$$\nu_{AB} = (h-1)\times(q-1), \quad \nu_{rez} = hq(n-1).$$

Cu aceste cantităţi se completează tabloul analizei dispersionale bifactoriale, tabelul 9.4, care permite testarea ipotezelor privind semnificaţia contribuţiei factorilor la explicarea variaţiei variabilei dependente. Construcţia tabloului este

similară analizei dispersionale unifactoriale, coloana mediilor sumelor de pătrate fiind obţinută prin raportarea sumelor de pătrate la numărul corespunzător de grade de libertate.

Tabelul 9.4. *Tabloul analizei dispersionale (ANOVA) bifactoriale*

Sursa de variaţie	Suma pătratelor	Grade de libertate	Media pătratelor	F
Între grupuri				
Factorul A	SP_A	ν_A	$s_A^2 = SP_A/\nu_A$	$F_A = s_A^2/s_{rez}^2$
Factorul B	SP_B	ν_B	$s_B^2 = SP_B/\nu_B$	$F_B = s_B^2/s_{rez}^2$
Interacţiunea AB	SP_{AB}	ν_{AB}	$s_{AB}^2 = SP_{AB}/\nu_{AB}$	$F_{AB} = s_{AB}^2/s_{rez}^2$
În interiorul grupurilor	SP_{rez}	ν_{rez}	$s_{rez}^2 = SP_{rez}/\nu_{rez}$	
Globală	SP_g	ν_g		

Ca şi în cazul analizei dispersionale unifactoriale, verificarea semnificaţiei factorilor se face prin intermediul unor teste F comparând variaţiile explicate cu variaţia reziduală. De aici coloana ultimă din tablou, care conţine cele trei statistici necesare testelor F, obţinute ca rapoarte având la numitor s_{rez}^2.

Ipotezele testate. Sunt trei seturi de ipoteze, corespunzătoare celor doi factori şi interacţiunii lor. Le vom nota după factorii implicaţi:

(AB)
- H_0 : interacţiunea factorilor A şi B nu este semnificativă (efectul lui A nu se modifică în funcţie de nivelul lui B)
- H_1 : interacţiunea factorilor A şi B este semnificativă (efectul lui A depinde de nivelul factorului B)

(A)
- H_0 : factorul A nu este semnificativ (clasificarea după linii nu este semnificativă sau mediile liniilor nu diferă semnificativ)
- H_1 : factorul A este semnificativ (clasificarea după linii este semnificativă, mediile liniilor diferă în mod semnificativ)

(B)
- H_0 : factorul B nu este semnificativ (clasificarea după coloane nu este semnificativă sau mediile coloanelor nu diferă semnificativ)
- H_1 : factorul B este semnificativ (clasificarea după coloane este semnificativă, mediile coloanelor diferă în mod semnificativ)

Decizia. Deoarece, acceptând ipoteza nulă a fiecărui test, se demonstrează că statistica F calculată în tabloul ANOVA urmează o repartiţie F având drept grade de libertate numerele asociate numărătorului şi numitorului, se va respinge ipoteza nulă H_0 în favoarea ipotezei alternative dacă valoarea F calculată, corespunzătoare

testului, citită în tabloul ANOVA, este mai mare sau egală cu quantila respectivă a repartiției Fisher-Snedecor.

Testul	Regula de respingere a ipotezei nule
(AB)	$F_{AB} \geq F_{1-\alpha;\nu_{AB};\nu_{rez}}$
(A)	$F_A \geq F_{1-\alpha;\nu_A;\nu_{rez}}$
(B)	$F_B \geq F_{1-\alpha;\nu_B;\nu_{rez}}$

9.2.2. Modele ale analizei dispersionale bifactoriale

După modul de obținere a nivelurilor factorilor, se pot imagina mai multe situații, după un raționament similar celui efectuat la analiza dispersională unifactorială:

1. Celulele sunt obținute la întâmplare dintr-o mulțime (populație) de celule posibile. Este ca și cum atât nivelurile factorului A, cât și nivelurile factorului B sunt mult mai numeroase decât cele considerate și dintre acestea s-au extras la întâmplare (într-un sondaj) nivelurile pentru care s-au efectuat determinări. Situația este desemnată ca fiind un **model cu efecte întâmplătoare**. Diferențele dintre celule sunt de tip aleatoriu, datorate, cel puțin parțial, sondajului.
2. Celulele considerate alcătuiesc întreaga mulțime de celule posibile. Factorii A și B nu au alte gradații în afara celor prezente. Modelul este denumit **model cu efecte sistematice**. Diferențele dintre celule sunt sistematice, datorate clasificării după cei doi factori.
3. Celulele s-au obținut sistematic după un factor (care apare cu toate nivelurile sale) și aleatoriu după celălalt factor (dintre nivelurile căruia apar doar o parte, extrase la întâmplare). Situația reprezintă **modelul cu efecte mixte**. Abaterile dintre celule sunt sistematice după un factor și întâmplătoare după celălalt.

Din punct de vedere matematic, se poate considera că valorile variabilei dependente X se obțin după o expresie liniară care leagă efectele factorilor de erorile nesistematice:

$$x_{ijk} = \mu + \alpha_i + \beta_j + \delta_{ij} + \varepsilon_{ijk}, \text{ pentru orice } i,j,k \text{ cu}$$
$$1 \leq i \leq h,\ 1 \leq j \leq q,\ 1 \leq k \leq n,$$

unde, pentru orice i, j, k în limitele admise, α_i reprezintă efectul factorului A; β_j reprezintă efectul factorului B; δ_{ij} reprezintă efectul interacțiunii factorilor A și B; ε_{ijk} este eroarea reziduală (datorată sondajului din cadrul subpopulației, datorată altor factori etc.), iar μ este media populației de bază (pentru care grupurile sunt subpopulații).

Asupra entităților introduse se presupune că sunt satisfăcute ipotezele:

- valorile medii ale tuturor efectelor și interacțiunilor sunt nule, adică:

$$\sum_i \alpha_i = \sum_j \beta_j = \sum_i \delta_{ij} = \sum_j \delta_{ij} = 0;$$

- erorile ε_{ijk} sunt v.a. independente, repartizate normal cu media zero și aceeași dispersie:

$$\varepsilon_{ijk} \sim N(0,\sigma^2),\ (\forall) i,j,k;$$

- valorile x_{ijk} sunt independente și repartizate normal (ipoteză necesară pentru determinarea intervalelor de încredere).

Trebuie să observăm faptul că ipotezele nu sunt mai restrictive decât cele presupuse în mod obișnuit pentru cantități asimilate unor erori experimentale.

Considerarea tipului de model al analizei aduce următoarele precizări:

1. în cazul modelului cu efecte sistematice, μ, α_i, β_j, δ_{ij} și σ^2 sunt mărimi fixe, necunoscute;
2. în cazul modelului cu efecte întâmplătoare, α_i, β_j și δ_{ij} sunt mărimi aleatorii cu dispersiile $\sigma^2_\alpha, \sigma^2_\beta, \sigma^2_{\alpha\beta}$; mărimile parametrilor μ, σ^2, $\sigma^2_\alpha, \sigma^2_\beta, \sigma^2_{\alpha\beta}$ sunt necunoscute;
3. în cazul modelului cu efecte mixte, presupunând că factorul A este cel întâmplător, μ și β_j sunt mărimi fixe (nealeatorii), iar α_i și δ_{ij} sunt mărimi aleatorii; $\mu, \beta_j, \sigma^2_\alpha, \sigma^2_{\alpha\beta}$ și σ^2 sunt parametri necunoscuți.

9.2.2.1. Modelul cu efecte sistematice[54]

Particularizând ipotezele generale testate, se poate considera că ipotezele nule sunt în acest model:

(AB) $H_0 : \delta_{ij} = 0$, pentru orice $i = 1, 2, \ldots, h; j = 1, 2, \ldots, q$;
(A) $H_0 : \alpha_i = 0$, pentru orice $i = 1, 2, \ldots, h$;
(B) $\beta_j = 0$, pentru orice $j = 1, 2, \ldots, q$.

Se va începe cu verificarea ipotezei privind absența interacțiunii factorilor – testul (AB). În funcție de rezultatul acestui test se modifică acțiunea ulterioară.

Nu se respinge ipoteza nulă a testului (AB). Acesta este cazul când se acceptă ipoteza că între cei doi factori nu există interacțiune, deci se poate considera că factorul A (sau B) nu își modifică influența în funcție de nivelurile celuilalt. În acest caz, se pot verifica ipotezele privind semnificația factorilor A și B.

54. Pentru celelalte modele ale analizei dispersionale bifactoriale se va vedea, de exemplu, N. Rancu și L. Tövissi, *Statistica matematică cu aplicații...*, Editura Academiei, București, 1963, sau N. Cojocaru; V. Clocotici; N. Dobră; *Metode statistice aplicate...*, Editura Tehnică, București, 1986 (vezi bibliografia).

Observaţie. Unii autori recomandă ca, în acest caz, linia din tabloul ANOVA care corespunde interacţiunii factorilor să fie adunată la linia sursei reziduale de variaţie, întrucât varianţa respectivă nu mai este explicată de cei doi factori.

Intervalele de încredere pentru parametrii necunoscuţi se obţin după formulele generale, cu menţiunea că dispersia reziduală este estimaţia pentru σ^2:

– media globală:

$$\bar{x} - t_{1-\alpha/2;\nu_{rez}} \frac{s^2_{rez}}{\sqrt{hqn}} \le \mu \le \bar{x} + t_{1-\alpha/2;\nu_{rez}} \frac{s^2_{rez}}{\sqrt{hqn}};$$

– media liniei i, $i = 1, 2, \ldots, h$:

$$\bar{x}_{i*} - t_{1-\alpha/2;\nu_{rez}} \frac{s^2_{rez}}{\sqrt{qn}} \le \mu_{i*} \le \bar{x}_{i*} + t_{1-\alpha/2;\nu_{rez}} \frac{s^2_{rez}}{\sqrt{qn}};$$

– media coloanei j, $j = 1, 2, \ldots, q$:

$$\bar{x}_{*j} - t_{1-\alpha/2;\nu_{rez}} \frac{s^2_{rez}}{\sqrt{hn}} \le \mu_{*j} \le \bar{x}_{*j} + t_{1-\alpha/2;\nu_{rez}} \frac{s^2_{rez}}{\sqrt{hn}};$$

– media celulei (i, j), $i = 1, 2, \ldots, h$; $j = 1, 2, \ldots, q$:

$$\bar{x}_{ij} - t_{1-\alpha/2;\nu_{rez}} \frac{s^2_{rez}}{\sqrt{n}} \le \mu_{ij} \le \bar{x}_{ij} + t_{1-\alpha/2;\nu_{rez}} \frac{s^2_{rez}}{\sqrt{n}};$$

– efectul α_i, $i = 1, 2, \ldots, h$:

$$(\bar{x}_{i*} - \bar{x}) - t_{1-\alpha/2;\nu_{rez}} \frac{s^2_{rez}\sqrt{h-1}}{\sqrt{hqn}} \le \alpha_i \le (\bar{x}_{i*} - \bar{x}) + t_{1-\alpha/2;\nu_{rez}} \frac{s^2_{rez}\sqrt{h-1}}{\sqrt{hqn}};$$

– efectul β_j, $j = 1, 2, \ldots, q$:

$$(\bar{x}_{*j} - \bar{x}) - t_{1-\alpha/2;\nu_{rez}} \frac{s^2_{rez}\sqrt{q-1}}{\sqrt{hqn}} \le \beta_j \le (\bar{x}_{*j} - \bar{x}) + t_{1-\alpha/2;\nu_{rez}} \frac{s^2_{rez}\sqrt{q-1}}{\sqrt{hqn}}.$$

Se respinge ipoteza nulă a testului (AB), deci se acceptă ipoteza că între factorii A şi B există o interacţiune semnificativă. În această situaţie, metoda analizei dispersionale nu mai conduce la rezultate valide, astfel încât se va recurge la alte metode, cum ar fi analiza de corelaţie sau cea de regresie.

9.3. Teste privind egalitatea dispersiilor unor populaţii normale

Deşi egalitatea unui număr de k dispersii apare la verificarea ipotezelor analizei dispersionale unifactoriale, problema prezintă interes şi în sine. A decide dacă împrăştierea valorilor se menţine constantă constituie o procedură importantă, complementară verificării egalităţii dintre medii, atunci când ne interesează mai mult de două (sub)populaţii. Pentru cazul a două populaţii, reamintim că se utilizează un test F. Problema poate fi desemnată ca reprezentând verificarea omogenităţii dispersiilor etc.

Pentru fixarea ideilor, considerăm un număr de k variabile aleatorii, X_i, $i = 1, 2, \ldots, k$, $(k > 2)$, repartizate normal $N(\mu_i; \sigma_i^2)$, atât mediile, cât şi dispersiile fiind necunoscute. Dispunem de dispersiile de sondaj $s_1^2, s_2^2, \ldots, s_k^2$, determinate din eşantioane de volume $n_1, n_2, \ldots, n_k$. Trebuie ca, pe baza datelor de sondaj, să se verifice ipoteza omogenităţii dispersiilor. Se consideră testul statistic cu ipotezele:

$$\begin{cases} H_0: \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2 \\ H_1: \text{există } i, j \text{ pentru care } \sigma_i^2 \neq \sigma_j^2 \text{ (există cel puţin două dispersii diferite)} \end{cases}$$

9.3.1. Testul Bartlett

Acest test este utilizat îndeosebi atunci când volumele eşantioanelor sunt diferite între ele.

Statistica testului este:

$$B = \frac{1}{c}\left[\sum_{i=1}^{k} \nu_i \ln s^2 - \sum_{i=1}^{k} \nu_i \ln s_i^2\right],$$

unde:

$$\nu_i = n_i - 1, \quad i = 1, 2, \ldots, k,$$

$$\nu = \sum_{i=1}^{k} \nu_i = \sum_{i=1}^{k} n_i - k, \quad s^2 = \frac{1}{\nu}\sum_{i=1}^{k} \nu_i s_i^2,$$

$$c = 1 + \frac{1}{3(k-1)}\left[\sum_{i=1}^{k} \frac{1}{\nu_i} - \frac{1}{\nu}\right].$$

Decizia în test se bazează pe faptul că, în condiţiile ipotezei nule, cantitatea B este repartizată χ^2 cu $k - 1$ grade de libertate.

Prin urmare, după logica generală a unui test bazat pe o repartiție cunoscută, se respinge ipoteza H_0 dacă valoarea B calculată este mai mare sau egală cu quantila corespunzătoare nivelului de semnificație fixat:

$$B \geq \chi^2_{1-\alpha;\, k-1}.$$

În această situație, se poate considera că dispersiile nu sunt omogene. Pentru a decide care două dispersii sunt diferite, se vor efectua teste F, comparând dispersiile două câte două.

Dacă nu este îndeplinită condiția de respingere a ipotezei nule, adică statistica B calculată este mai mică decât quantila $\chi^2_{1-\alpha;\, k-1}$, atunci se va accepta că dispersiile sunt omogene sau, în sensul raționamentului statistic, că datele de sondaj disponibile nu permit respingerea omogenității dispersiilor.

9.3.2. Testele Hartley şi Cochran

Pentru eşantioane de volume egale se pot utiliza aceste teste, ale căror statistici sunt mult mai simplu de calculat (şi testul Bartlett devine ceva mai simplu de aplicat pentru eşantioane egale).

Statisticile testelor

Hartley – raportul dintre dispersia de sondaj maximă şi dispersia de sondaj minimă:

$$F_{max} = \frac{\max\limits_i s_i^2}{\min\limits_i s_i^2}.$$

Cochran – raportul dintre dispersia de sondaj maximă şi suma tuturor dispersiilor de sondaj:

$$Q_{max} = \frac{\max\limits_i s_i^2}{\sum\limits_{i=1}^{k} s_i^2}.$$

Decizia. Ipoteza nulă a egalității dispersiilor este respinsă în cazul în care statistica este mai mare sau egală cu valoarea critică ce corespunde testului, volumului n al eşantioanelor (acelaşi la toate), numărului k de dispersii comparate şi nivelului α de semnificație fixat:

$$F_{max} \geq F_{1-\alpha;\, \nu_1=k;\, \nu_2=n-1} \text{ sau}$$

$$Q_{max} \geq Q_{1-\alpha;\, \nu_1=k;\, \nu_2=n-1},$$

unde primul număr de grade de libertate este numărul de grupuri comparate, iar al doilea număr de grade de libertate este volumul eşantionului mai puțin unu.

Valorile critice pentru testul Hartley sunt date în tabelul 9.5, iar pentru testul Cochran în tabelul 9.6. Valorile sunt date doar pentru un număr rezonabil de dispersii comparate şi doar pentru $\alpha = 0{,}05$.

Interpretarea respingerii sau tolerării ipotezei nule este aceeaşi ca la testul Bartlett.

Tabelul 9.5. *Valorile critice $F_{\alpha=0,05;\ \nu1=k;\ \nu2=n-1}$ pentru aplicarea testului Hartley*

n–1 \ k	3	4	5	6	7	8
2	87,50	142,00	202,00	266,00	333,00	403,00
3	27,80	39,20	50,70	62,00	72,90	83,50
4	15,50	20,60	25,20	29,50	33,60	37,50
5	10,80	13,70	16,30	18,70	20,80	22,90
6	8,38	10,40	12,10	13,70	15,00	16,30
7	6,94	8,44	9,70	10,80	11,80	12,70
8	6,00	7,18	8,12	9,03	90,78	10,50
9	5,34	6,31	7,11	7,80	8,41	8,95
10	4,85	5,67	6,34	6,92	7,42	7,87
12	4,16	4,79	5,30	5,72	6,09	6,42
15	3,54	4,01	4,37	4,68	4,95	5,19
20	2,95	3,29	3,54	3,76	3,94	4,10
30	2,40	2,61	2,78	2,91	3,02	3,12
60	1,85	1,96	2,04	2,11	2,17	2,22

Tabelul 9.6. *Valorile critice $Q_{\alpha=0,05;\ \nu1=k;\ \nu2=n-1}$ pentru aplicarea testului Cochran*

n-1 \ k	3	4	5	6	7	8
1	0,9669	0,9065	0,8412	0,7808	0,7271	0,6798
2	0,8709	0,7679	0,6838	0,6161	0,5612	0,5157
3	0,7977	0,6841	0,5981	0,5321	0,4800	0,4307
4	0,7457	0,6287	0,5440	0,4803	0,4307	0,3910
5	0,7074	0,5895	0,5063	0,4447	0,3974	0,3595
6	0,6771	0,5598	0,4783	0,4184	0,3726	0,3362
7	0,6530	0,5365	0,4564	0,3980	0,3535	0,3185
8	0,6333	0,5175	0,4387	0,3817	0,3384	0,3043
9	0,6167	0,5017	0,4241	0,3682	0,3259	0,2926
10	0,6025	0,4884	0,4118	0,3568	0,3154	0,2829
16	0,5466	0,4366	0,3645	0,3135	0,2756	0,2462
36	0,4748	0,3720	0,3066	0,2612	0,2278	0,2022

9.4. Aplicație

Deoarece problematica acestui capitol este mult mai unitară în raport cu capitolele precedente, exemplele sunt mai sugestive dacă se referă la întregul capitol. Fiecare exemplu se vrea și un model de analiză a datelor, chiar dacă, în bună parte, ele sunt didactice: pentru a fi mai complexe și a oferi posibilitatea de a comenta mai multe situații ce pot să apară în practică, unele valori au fost modificate, iar altele au fost eliminate.

Exemplu. Într-un studiu privind însușirea de către elevi a unei limbi străine s-au obținut următoarele date experimentale. S-au considerat grupuri de elevi după locul de reședință și după calificarea profesorului care predă disciplina respectivă. Notele obținute de elevii testați sunt cuprinse în tabelul 9.7.

Tabelul 9.7. *Datele de sondaj pentru exemplu*

Mediu / Profesor	Urban	Suburban	Rural
Grad I	9	9	7
	8	8	6
	10	8	9
	8	9	7
	7	8	8
Grad II	9	6	7
	8	8	9
	9	5	5
	6	6	7
	9	7	4
Definitivat	7	7	6
	5	7	7
	8	4	6
	6	7	7
	7	6	5
Stagiar	8	8	5
	8	7	7
	6	5	8
	6	7	5
	7	6	7

Întrebările la care vom încerca să răspundem sunt:

- limba străină care face obiectul cercetării este la fel de bine însușită atât la oraș. cât și la sat?
- însușirea limbii străine depinde de gradul de calificare profesională a profesorului care o predă?
- profesorii au aceeași eficiență, indiferent de mediul în care își desfășoară activitatea didactică?

Reformularea acestor întrebări în termenii prelucrărilor statistice duce la întrebări de genul:

1. Există diferențe între subpopulațiile (grupurile) de elevi determinate de mediu?
2. Există diferențe între subpopulațiile (grupurile) de elevi determinate de calificarea profesorului?
3. Există diferențe între subpopulațiile (grupurile) determinate de mediu *și* de profesor?

Sunt posibile mai multe moduri de abordare a acestor probleme. În capitolul de față vom utiliza, în principal pentru exemplificare, metoda analizei dispersionale.

Problema 1. Calculul mediilor și dispersiilor pentru grupurile determinate de mediu, deci pentru coloanele tabelului 9.7, dă valorile:

	Urban	Suburban	Rural
Medii	7,55	6,90	6,60
Dispersii	1,73	1,78	1,83
N_i	$n_1 = 20$	$n_2 = 20$	$n_3 = 20$

Aplicând testul Hartley pentru verificarea omogenității dispersiilor, obținem:

$$F_{max} = \frac{\max s_i^2}{\min s_i^2} = \frac{s_3^2}{s_1^2} = \frac{1,83}{1,73} = 1,058.$$

Cum $\nu_1 = 3$ (numărul de grupuri) și $\nu_2 = 19$ (numărul de observații minus unu), în tabelul valorilor critice ale testului Hartley găsim $F_{0,95;3;20} = 2,95$ (cel mai apropiat de parametrii noștri). Valoarea calculată este mai mică decât valoarea critică; prin urmare, nu putem respinge ipoteza nulă a egalității dispersiilor la nivelul de semnificație 0,05. Vom accepta deci că dispersiile subpopulațiilor sunt egale (diferențele care apar, naturale în orice sondaj, nu susțin ipoteza unor dispersii diferite).

Compararea grupurilor presupune atunci compararea mediilor lor. Putem aplica în acest sens o succesiune de teste Student. Obținem următoarele valori ale statisticilor calculate (teste t presupunând dispersiile egale):

Grupurile comparate	t
urban – suburban	1,55
urban – rural	2,25
suburban – rural	0,76

Urmând procedura de decizie a unui test t bilateral, la nivelul de semnificație $\alpha = 0,05$, găsim în tabelul repartiției Student $t_{0,97540} = 2,021$. Numărul exact al gradelor de libertate este 38; de aceea, am luat valoarea cea mai apropiată (40) dintre cele listate de obicei în tabele (valoarea exactă este $t_{0,97538} = 2,024$). Cele trei teste arată atunci că singura diferență care pare semnificativă este aceea dintre urban și rural (doar 2,25 > 2,021). În acest caz, testul unilateral permite chiar să se accepte ipoteza unilaterală că media din urban este mai mare decât media din rural.

Argumentația de la începutul acestui capitol arată însă că utilizarea unei asemenea proceduri de verificări multiple mărește riscul apariției erorilor de speța întâi. Cu alte cuvinte, este de așteptat să crească probabilitatea ca, într-un număr de comparații două câte două, să apară în mod eronat diferențe semnificative. Din acest motiv este corect să se analizeze mai întâi situația globală. Dezideratul respectiv se realizează aplicând analiza dispersională unifactorială. Tabloul ANOVA obținut arată astfel:

Sursa variației	Suma pătratelor	Grade de libertate	Media pătratelor	F
Între grupuri (externă)	9,433	2	4,717	2,647
În grupuri (internă)	101,550	57	1,782	
Globală	110,983	59		

La nivelul $\alpha = 0,05$ obținem $F_{0,95;2;57} = 3,159$. Valoarea calculată, 2,647, este mai mică decât valoarea critică (la nivelul fixat de 0,05). Decizia în test vizează, prin urmare, faptul că nu putem respinge ipoteza nulă a egalității mediilor celor trei (sub)populații de elevi. Datele de sondaj, dacă neglijăm total clasificarea după profesori, nu susțin existența diferențelor între elevii din cele trei medii avute în vedere.

Problema 2. Considerând grupurile determinate de gradul didactic al profesorilor, deci pentru liniile (de celule) din tabelul 9.7, obținem următoarele valori ale mediilor și dispersiilor de sondaj:

	Grad I	Grad II	Definitivat	Stagiar
Medii	8,067	7,000	6,333	6,667
Dispersii	1,067	2,714	1,095	1,238
N_i	15	15	15	15

Verificarea omogenității dispersiilor prin testul Hartley va da:

$$F_{max} = \frac{\max s_i^2}{\min s_i^2} = \frac{s_2^2}{s_1^2} = \frac{2,714}{1,067} = 2,544.$$

Cum $\nu_1 = 4$ (numărul de grupuri) și $\nu_2 = 14$ (numărul de observații minus unu), în tabelul valorilor critice ale testului Hartley găsim $F_{0,95;4;14} = 4,01$ (cel mai apropiat de parametrii noștri). Valoarea calculată este mai mică decât valoarea critică; prin urmare, nu putem respinge ipoteza nulă a egalității dispersiilor, la nivelul de semnificație $\alpha = 0,05$. Vom accepta deci ideea că dispersiile subpopulațiilor sunt egale (diferențele care apar, naturale în orice sondaj, nu susțin ipoteza unor dispersii diferite).

Comparațiile între mediile grupurilor presupun 6 teste Student (sunt 6 perechi de grupuri). Dintre acestea se observă că media grupului „grad I" diferă de media grupului „grad II" (t = 2,125), de media grupului „definitivat" (t = 4,566) etc. Mediile grupurilor „definitivat" și „stagiar" nu diferă semnificativ (t = –0,845) etc. Deja, datorită numărului mare de comparații de câte două grupuri, situația de ansamblu devine mai puțin clară. Din acest motiv, pe lângă cele specificate la începutul capitolului, se impune utilizarea unei comparații globale. Tabloul ANOVA pentru analiza dispersională unifactorială este:

Sursa variației	Suma pătratelor	Grade de libertate	Media pătratelor	F
Între grupuri (externă)	25,383	3	8,461	5,534
În grupuri (internă)	85,600	56	1,529	
Globală	110,983	59		

La nivelul $\alpha = 0,05$ obținem $F_{0,95;3;56} = 2,77$. Valoarea calculată, 5,534, este mai mare decât valoarea critică (la nivelul fixat $\alpha = 0,05$). Rezultă că putem respinge ipoteza nulă a egalității mediilor celor patru (sub)populații de elevi. În acest caz, prezintă importanță recurgerea la comparațiile individuale ale mediilor (două câte două). Va rezulta că respingerea ipotezei globale a omogenității mediilor se datorează faptului că media grupului determinat de profesorii de gradul I este diferită de restul mediilor.

Problema 3. Pentru răspunsul la această întrebare nu mai putem, practic, recurge la procedura utilizată pentru primele două probleme. Plecând de la clasificarea încrucișată după categoriile nominale ale variabilelor *mediu* și *grad profesional*, rezultă 12 grupuri, ceea ce presupune $C_{12}^2 = 66$ comparații distincte între grupuri. Interpretarea celor 66 de rezultate devine foarte complicată. Din acest motiv, pe lângă celelalte argumente deja aduse, se impune utilizarea analizei dispersionale bifactoriale. Cei doi factori (variabilele independente) sunt variabilele *mediu* și *grad profesional*, variabila dependentă fiind *nota*.

Tabloul ANOVA bifactorial este:

Sursa de variație	Suma pătratelor	Grade de libertate	Media pătratelor	F
Între grupuri				
Factorul *mediu*	25,383	3	8,461	5,803
Factorul *grad*	9,433	2	4,717	3,235
Interacțiunea *mediu – grad*	6,167	6	1,028	0,705
În interiorul grupurilor	70	48	1,458	
Globală	110,983	59		

La pragul $\alpha = 0,05$ valorile critice ale testului F sunt $F_{0,95;3;48}=2,80$, $F_{0,95;2;48} = 3,19$ și $F_{0,95;6;48}=2,29$. Comparând valorile F calculate cu quantilele corespunzătoare ale repartiției Fisher, se ajunge la concluzia că factorii mediu și grad sunt semnificativi în explicarea diferențelor dintre note, în timp ce nu se poate susține semnificația interacțiunii dintre factorii semnificativi.

Comparând concluzia din problema 3 cu cea obținută în problema 1, se ajunge la o aparentă contradicție (în problema 1 factorul mediu nu era semnificativ). Contradicția se poate rezolva prin aceea că, dacă din varianța globală se separă partea explicată de factorul grad, atunci partea explicată de factorul mediu devine suficient de mare față de partea reziduală, astfel încât devine semnificativă. Se susține astfel ideea că o problemă multivariată trebuie să fie tratată printr-o tehnică specifică, orice abordare parțială reliefând doar anumite aspecte, diferite uneori de cele obținute prin metoda globală.

ANEXE

Anexa 1. Funcția de repartiție normală standard N(0,1)

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998

Anexa 2. Quantilele distribuției χ^2,
având probabilitatea $1-P = \alpha$ de a fi depășite

ν \ α	0,995	0,99	0,975	0,95	0,9	0,1	0,05	0,025	0,01	0,005	0,001
1	0,0000	0,0002	0,0010	0,0039	0,0158	2,706	3,841	5,024	6,635	7,879	10,83
2	0,0100	0,0201	0,0506	0,1026	0,2107	4,605	5,991	7,378	9,210	10,60	13,82
3	0,072	0,115	0,216	0,352	0,584	6,251	7,815	9,348	11,34	12,84	16,27
4	0,207	0,297	0,484	0,711	1,064	7,779	9,488	11,14	13,28	14,86	18,47
5	0,412	0,554	0,831	1,145	1,610	9,236	11,07	12,83	15,09	16,75	20,51
6	0,676	0,872	1,237	1,635	2,204	10,64	12,59	14,45	16,81	18,55	22,46
7	0,989	1,239	1,690	2,167	2,833	12,02	14,07	16,01	18,48	20,28	24,32
8	1,344	1,647	2,180	2,733	3,490	13,36	15,51	17,53	20,09	21,95	26,12
9	1,735	2,088	2,700	3,325	4,168	14,68	16,92	19,02	21,67	23,59	27,88
10	2,156	2,558	3,247	3,940	4,865	15,99	18,31	20,48	23,21	25,19	29,59
11	2,603	3,053	3,816	4,575	5,578	17,28	19,68	21,92	24,73	26,76	31,26
12	3,074	3,571	4,404	5,226	6,304	18,55	21,03	23,34	26,22	28,30	32,91
13	3,565	4,107	5,009	5,892	7,041	19,81	22,36	24,74	27,69	29,82	34,53
14	4,075	4,660	5,629	6,571	7,790	21,06	23,68	26,12	29,14	31,32	36,12
15	4,601	5,229	6,262	7,261	8,547	22,31	25,00	27,49	30,58	32,80	37,70
16	5,142	5,812	6,908	7,962	9,312	23,54	26,30	28,85	32,00	34,27	39,25
17	5,697	6,408	7,564	8,672	10,09	24,77	27,59	30,19	33,41	35,72	40,79
18	6,265	7,015	8,231	9,390	10,86	25,99	28,87	31,53	34,81	37,16	42,31
19	6,844	7,633	8,907	10,12	11,65	27,20	30,14	32,85	36,19	38,58	43,82
20	7,434	8,260	9,591	10,85	12,44	28,41	31,41	34,17	37,57	40,00	45,31
21	8,034	8,897	10,28	11,59	13,24	29,62	32,67	35,48	38,93	41,40	46,80
22	8,643	9,542	10,98	12,34	14,04	30,81	33,92	36,78	40,29	42,80	48,27
23	9,260	10,20	11,69	13,09	14,85	32,01	35,17	38,08	41,64	44,18	49,73
24	9,886	10,86	12,40	13,85	15,66	33,20	36,42	39,36	42,98	45,56	51,18
25	10,52	11,52	13,12	14,61	16,47	34,38	37,65	40,65	44,31	46,93	52,62
26	11,16	12,20	13,84	15,38	17,29	35,56	38,89	41,92	45,64	48,29	54,05
27	11,81	12,88	14,57	16,15	18,11	36,74	40,11	43,19	46,96	49,65	55,48
28	12,46	13,56	15,31	16,93	18,94	37,92	41,34	44,46	48,28	50,99	56,89
29	13,12	14,26	16,05	17,71	19,77	39,09	42,56	45,72	49,59	52,34	58,30
30	13,79	14,95	16,79	18,49	20,60	40,26	43,77	46,98	50,89	53,67	59,70

Anexa 3. Quantilele repartiției Student, $t_{\nu;1-\alpha}$, pentru care $P(t \leq t_{\nu;1-\alpha}) = 1-\alpha$

ν \ 1−α	0,9	0,95	0,975	0,99	0,995	0,999	0,9995
1	3,078	6,314	12,706	31,821	63,656	318,289	636,578
2	1,886	2,920	4,303	6,965	9,925	22,328	31,600
3	1,638	2,353	3,182	4,541	5,841	10,214	12,924
4	1,533	2,132	2,776	3,747	4,604	7,173	8,610
5	1,476	2,015	2,571	3,365	4,032	5,894	6,869
6	1,440	1,943	2,447	3,143	3,707	5,208	5,959
7	1,415	1,895	2,365	2,998	3,499	4,785	5,408
8	1,397	1,860	2,306	2,896	3,355	4,501	5,041
9	1,383	1,833	2,262	2,821	3,250	4,297	4,781
10	1,372	1,812	2,228	2,764	3,169	4,144	4,587
11	1,363	1,796	2,201	2,718	3,106	4,025	4,437
12	1,356	1,782	2,179	2,681	3,055	3,930	4,318
13	1,350	1,771	2,160	2,650	3,012	3,852	4,221
14	1,345	1,761	2,145	2,624	2,977	3,787	4,140
15	1,341	1,753	2,131	2,602	2,947	3,733	4,073
16	1,337	1,746	2,120	2,583	2,921	3,686	4,015
17	1,333	1,740	2,110	2,567	2,898	3,646	3,965
18	1,330	1,734	2,101	2,552	2,878	3,610	3,922
19	1,328	1,729	2,093	2,539	2,861	3,579	3,883
20	1,325	1,725	2,086	2,528	2,845	3,552	3,850
21	1,323	1,721	2,080	2,518	2,831	3,527	3,819
22	1,321	1,717	2,074	2,508	2,819	3,505	3,792
23	1,319	1,714	2,069	2,500	2,807	3,485	3,768
24	1,318	1,711	2,064	2,492	2,797	3,467	3,745
25	1,316	1,708	2,060	2,485	2,787	3,450	3,725
26	1,315	1,706	2,056	2,479	2,779	3,435	3,707
27	1,314	1,703	2,052	2,473	2,771	3,421	3,689
28	1,313	1,701	2,048	2,467	2,763	3,408	3,674
29	1,311	1,699	2,045	2,462	2,756	3,396	3,660
30	1,310	1,697	2,042	2,457	2,750	3,385	3,646
40	1,303	1,684	2,021	2,423	2,704	3,307	3,551
50	1,299	1,676	2,009	2,403	2,678	3,261	3,496
60	1,296	1,671	2,000	2,390	2,660	3,232	3,460
80	1,292	1,664	1,990	2,374	2,639	3,195	3,416
100	1,290	1,660	1,984	2,364	2,626	3,174	3,390
200	1,286	1,653	1,972	2,345	2,601	3,131	3,340

Anexa 4. Quantilele repartiției Fisher-Snedecor, $F_{v1,v2;0.95}$, pentru $\alpha = 0,95$

v2 \ v1	1	2	3	4	5	6	7	8	9	10	15
1	161	199	216	225	230	234	237	239	241	242	246
2	18,5	19,0	19,2	19,2	19,3	19,3	19,4	19,4	19,4	19,4	19,4
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,70
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,86
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,62
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	3,94
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,51
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,22
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,01
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,85
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,72
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,62
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,53
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,46
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,40
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,35
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,31
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,27
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,23
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,20
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,18
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,15
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,13
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,11
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,09
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,07
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,06
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,04
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,03
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,01
32	4,15	3,29	2,90	2,67	2,51	2,40	2,31	2,24	2,19	2,14	1,99
34	4,13	3,28	2,88	2,65	2,49	2,38	2,29	2,23	2,17	2,12	1,97
36	4,11	3,26	2,87	2,63	2,48	2,36	2,28	2,21	2,15	2,11	1,95
38	4,10	3,24	2,85	2,62	2,46	2,35	2,26	2,19	2,14	2,09	1,94
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	1,92

Anexa 4. Quantilele repartiției Fisher-Snedecor, $F_{v1,v2;0.95}$ (continuare)

v2 \ v1	20	25	30	35	40	45	50	60	80	100	200
1	248	249	250	251	251	251	252	252	253	253	254
2	19,45	19,46	19,46	19,47	19,47	19,47	19,48	19,48	19,48	19,49	19,49
3	8,66	8,63	8,62	8,60	8,59	8,59	8,58	8,57	8,56	8,55	8,54
4	5,80	5,77	5,75	5,73	5,72	5,71	5,70	5,69	5,67	5,66	5,65
5	4,56	4,52	4,50	4,48	4,46	4,45	4,44	4,43	4,41	4,41	4,39
6	3,87	3,83	3,81	3,79	3,77	3,76	3,75	3,74	3,72	3,71	3,69
7	3,44	3,40	3,38	3,36	3,34	3,33	3,32	3,30	3,29	3,27	3,25
8	3,15	3,11	3,08	3,06	3,04	3,03	3,02	3,01	2,99	2,97	2,95
9	2,94	2,89	2,86	2,84	2,83	2,81	2,80	2,79	2,77	2,76	2,73
10	2,77	2,73	2,70	2,68	2,66	2,65	2,64	2,62	2,60	2,59	2,56
11	2,65	2,60	2,57	2,55	2,53	2,52	2,51	2,49	2,47	2,46	2,43
12	2,54	2,50	2,47	2,44	2,43	2,41	2,40	2,38	2,36	2,35	2,32
13	2,46	2,41	2,38	2,36	2,34	2,33	2,31	2,30	2,27	2,26	2,23
14	2,39	2,34	2,31	2,28	2,27	2,25	2,24	2,22	2,20	2,19	2,16
15	2,33	2,28	2,25	2,22	2,20	2,19	2,18	2,16	2,14	2,12	2,10
16	2,28	2,23	2,19	2,17	2,15	2,14	2,12	2,11	2,08	2,07	2,04
17	2,23	2,18	2,15	2,12	2,10	2,09	2,08	2,06	2,03	2,02	1,99
18	2,19	2,14	2,11	2,08	2,06	2,05	2,04	2,02	1,99	1,98	1,95
19	2,16	2,11	2,07	2,05	2,03	2,01	2,00	1,98	1,96	1,94	1,91
20	2,12	2,07	2,04	2,01	1,99	1,98	1,97	1,95	1,92	1,91	1,88
21	2,10	2,05	2,01	1,98	1,96	1,95	1,94	1,92	1,89	1,88	1,84
22	2,07	2,02	1,98	1,96	1,94	1,92	1,91	1,89	1,86	1,85	1,82
23	2,05	2,00	1,96	1,93	1,91	1,90	1,88	1,86	1,84	1,82	1,79
24	2,03	1,97	1,94	1,91	1,89	1,88	1,86	1,84	1,82	1,80	1,77
25	2,01	1,96	1,92	1,89	1,87	1,86	1,84	1,82	1,80	1,78	1,75
26	1,99	1,94	1,90	1,87	1,85	1,84	1,82	1,80	1,78	1,76	1,73
27	1,97	1,92	1,88	1,86	1,84	1,82	1,81	1,79	1,76	1,74	1,71
28	1,96	1,91	1,87	1,84	1,82	1,80	1,79	1,77	1,74	1,73	1,69
29	1,94	1,89	1,85	1,83	1,81	1,79	1,77	1,75	1,73	1,71	1,67
30	1,93	1,88	1,84	1,81	1,79	1,77	1,76	1,74	1,71	1,70	1,66
32	1,91	1,85	1,82	1,79	1,77	1,75	1,74	1,71	1,69	1,67	1,63
34	1,89	1,83	1,80	1,77	1,75	1,73	1,71	1,69	1,66	1,65	1,61
36	1,87	1,81	1,78	1,75	1,73	1,71	1,69	1,67	1,64	1,62	1,59
38	1,85	1,80	1,76	1,73	1,71	1,69	1,68	1,65	1,62	1,61	1,57
40	1,84	1,78	1,74	1,72	1,69	1,67	1,66	1,64	1,61	1,59	1,55

Anexa 4. Quantilele repartiției Fisher-Snedecor, $F_{v1,v2;0.95}$ (continuare)

v2 \ v1	1	2	3	4	5	6	7	8	9	10	15
42	4,07	3,22	2,83	2,59	2,44	2,32	2,24	2,17	2,11	2,06	1,91
44	4,06	3,21	2,82	2,58	2,43	2,31	2,23	2,16	2,10	2,05	1,90
46	4,05	3,20	2,81	2,57	2,42	2,30	2,22	2,15	2,09	2,04	1,89
48	4,04	3,19	2,80	2,57	2,41	2,29	2,21	2,14	2,08	2,03	1,88
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,03	1,87
52	4,03	3,18	2,78	2,55	2,39	2,28	2,19	2,12	2,07	2,02	1,86
54	4,02	3,17	2,78	2,54	2,39	2,27	2,18	2,12	2,06	2,01	1,86
56	4,01	3,16	2,77	2,54	2,38	2,27	2,18	2,11	2,05	2,00	1,85
58	4,01	3,16	2,76	2,53	2,37	2,26	2,17	2,10	2,05	2,00	1,84
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,84
62	4,00	3,15	2,75	2,52	2,36	2,25	2,16	2,09	2,03	1,99	1,83
64	3,99	3,14	2,75	2,52	2,36	2,24	2,16	2,09	2,03	1,98	1,83
66	3,99	3,14	2,74	2,51	2,35	2,24	2,15	2,08	2,03	1,98	1,82
68	3,98	3,13	2,74	2,51	2,35	2,24	2,15	2,08	2,02	1,97	1,82
70	3,98	3,13	2,74	2,50	2,35	2,23	2,14	2,07	2,02	1,97	1,81
75	3,97	3,12	2,73	2,49	2,34	2,22	2,13	2,06	2,01	1,96	1,80
80	3,96	3,11	2,72	2,49	2,33	2,21	2,13	2,06	2,00	1,95	1,79
85	3,95	3,10	2,71	2,48	2,32	2,21	2,12	2,05	1,99	1,94	1,79
90	3,95	3,10	2,71	2,47	2,32	2,20	2,11	2,04	1,99	1,94	1,78
95	3,94	3,09	2,70	2,47	2,31	2,20	2,11	2,04	1,98	1,93	1,77
100	3,94	3,09	2,70	2,46	2,31	2,19	2,10	2,03	1,97	1,93	1,77
125	3,92	3,07	2,68	2,44	2,29	2,17	2,08	2,01	1,96	1,91	1,75
150	3,90	3,06	2,66	2,43	2,27	2,16	2,07	2,00	1,94	1,89	1,73
175	3,90	3,05	2,66	2,42	2,27	2,15	2,06	1,99	1,93	1,89	1,72
200	3,89	3,04	2,65	2,42	2,26	2,14	2,06	1,98	1,93	1,88	1,72
250	3,88	3,03	2,64	2,41	2,25	2,13	2,05	1,98	1,92	1,87	1,71
300	3,87	3,03	2,63	2,40	2,24	2,13	2,04	1,97	1,91	1,86	1,70
500	3,86	3,01	2,62	2,39	2,23	2,12	2,03	1,96	1,90	1,85	1,69
1000	3,85	3,00	2,61	2,38	2,22	2,11	2,02	1,95	1,89	1,84	1,68

Anexa 4. Quantilele repartiției Fisher-Snedecor, $F_{v1,v2;0.95}$ (continuare)

v2 \ v1	20	25	30	35	40	45	50	60	80	100	200
42	1,83	1,77	1,73	1,70	1,68	1,66	1,65	1,62	1,59	1,57	1,53
44	1,81	1,76	1,72	1,69	1,67	1,65	1,63	1,61	1,58	1,56	1,52
46	1,80	1,75	1,71	1,68	1,65	1,64	1,62	1,60	1,57	1,55	1,51
48	1,79	1,74	1,70	1,67	1,64	1,62	1,61	1,59	1,56	1,54	1,49
50	1,78	1,73	1,69	1,66	1,63	1,61	1,60	1,58	1,54	1,52	1,48
52	1,78	1,72	1,68	1,65	1,62	1,61	1,59	1,57	1,53	1,51	1,47
54	1,77	1,71	1,67	1,64	1,62	1,60	1,58	1,56	1,53	1,51	1,46
56	1,76	1,70	1,66	1,63	1,61	1,59	1,57	1,55	1,52	1,50	1,45
58	1,75	1,70	1,66	1,63	1,60	1,58	1,57	1,54	1,51	1,49	1,45
60	1,75	1,69	1,65	1,62	1,59	1,57	1,56	1,53	1,50	1,48	1,44
62	1,74	1,68	1,64	1,61	1,59	1,57	1,55	1,53	1,49	1,47	1,43
64	1,74	1,68	1,64	1,61	1,58	1,56	1,55	1,52	1,49	1,47	1,42
66	1,73	1,67	1,63	1,60	1,58	1,56	1,54	1,52	1,48	1,46	1,42
68	1,73	1,67	1,63	1,60	1,57	1,55	1,54	1,51	1,48	1,46	1,41
70	1,72	1,66	1,62	1,59	1,57	1,55	1,53	1,50	1,47	1,45	1,40
75	1,71	1,65	1,61	1,58	1,55	1,53	1,52	1,49	1,46	1,44	1,39
80	1,70	1,64	1,60	1,57	1,54	1,52	1,51	1,48	1,45	1,43	1,38
85	1,70	1,64	1,59	1,56	1,54	1,52	1,50	1,47	1,44	1,42	1,37
90	1,69	1,63	1,59	1,55	1,53	1,51	1,49	1,46	1,43	1,41	1,36
95	1,68	1,62	1,58	1,55	1,52	1,50	1,48	1,46	1,42	1,40	1,35
100	1,68	1,62	1,57	1,54	1,52	1,49	1,48	1,45	1,41	1,39	1,34
125	1,66	1,59	1,55	1,52	1,49	1,47	1,45	1,42	1,39	1,36	1,31
150	1,64	1,58	1,54	1,50	1,48	1,45	1,44	1,41	1,37	1,34	1,29
175	1,63	1,57	1,52	1,49	1,46	1,44	1,42	1,39	1,36	1,33	1,27
200	1,62	1,56	1,52	1,48	1,46	1,43	1,41	1,39	1,35	1,32	1,26
250	1,61	1,55	1,50	1,47	1,44	1,42	1,40	1,37	1,33	1,31	1,25
300	1,61	1,54	1,50	1,46	1,43	1,41	1,39	1,36	1,32	1,30	1,23
500	1,59	1,53	1,48	1,45	1,42	1,40	1,38	1,35	1,30	1,28	1,21
1000	1,58	1,52	1,47	1,43	1,41	1,38	1,36	1,33	1,29	1,26	1,19

Bibliografie

***, *Grand dictionnaire de la psychologie*, Larousse, Paris, 1992.

Andrei, T., Stancu, S., *Statistica. Teorie şi aplicaţii*, Editura All, Bucureşti, 1995.

Barais, Annick Weil (coord.), *Les méthodes en psychologie*, Collection Grand Amphi, Breal, Rosny, 1997.

Bennet, A.C.; Franklin, L.N., *Statistical Analysis in Chemistry and the Chemical Industry*, Wiley, New York, 1954.

Bernier, J. J., Pietrulewicz, B., *La psychométrie*, Gaëtan Morin (éd.), Montréal, Paris, Casablanca, 1997.

Bryant, E.C., *Statistical Analysis*, McGraw-Hill, New York, 1960.

Ciucu, G., Craiu, V., *Teoria estimaţiei şi verificarea ipotezelor statistice*, Editura Didactică şi Pedagogică, Bucureşti, 1969.

Cochran, W.G., *Stichprobenverfahren*, W. de Gruyter, New York, 1972.

Cojocaru, N.; Clocotici, V.; Dobră, N., *Metode statistice aplicate în industria textilă*, Editura Tehnică, Bucureşti, 1986.

Craiu, V., *Verificarea ipotezelor statistice*, Editura Didactică şi Pedagogică, Bucureşti, 1972.

Cramer, H., *Mathematical Methods of Statistics*, Princeton University Press, 1961.

Dickes, Paul; Tournois, Jocelyne; Flieller, André; Kop, Jean-Luc, *La psychométrie. Théorie et méthode de la mesure*, Presses Universitaires de France, Paris, 1994.

Ghiglione, R.; Richard, J.F. (sous la direction de), *Cours de Psychologie*, Dunod, Paris, 1995.

Gueguen, Nicolas, *Manuel de statistiques pour psychologues*, Dunod, Paris, 1997.

Havârneanu, Cornel, *Cunoaşterea psihologică a persoanei*, Editura Polirom, Iaşi, 2000.

Hersen, M.; Kazdin, A.E.; Bellack, A.E. (eds.), *The Clinical Psychology Handbook*, Pergamon Press, New York, Oxford, 1991.

Hofstätter, Peter R., *Psychologie*, Fischer Taschenbuch Verlag, Frankfurt-pe-Main, 1971.

Horst, P., *Messung und Vorhersage*, Verlag Juliusz Beltz, Weinheim, Berlin, Basel, 1971.

Iosifescu, M.; Moineagu, C.; Trabici, V.; Ursianu, E., *Mică enciclopedie de statistică*, Editura Ştiinţifică şi Enciclopedică, Bucureşti, 1985.

Isaic, Al., *Statistica teoretică economică*, Editura Tehnică, Chişinău, 1994.

Kurtz, Norman R., *Introduction to Social Statistics*, McGraw-Hill Book Company, New York, 1983.

Landau, Erika, *Psychologie der Kreativität*, Ernst Reinhardt Verlag, München, Basel, 1969.

Lienert, Gustav A., *Testaufbau und Testanalyse*, Verlag Juliusz Beltz, Weinheim, Berlin, 1967.

Magnusson, David, *Testtheorie*, Verlag Franz Deuticke, Viena, 1975.

Miclea, Mircea, *Psihologie cognitivă*, Casa de editură „Gloria" S.R.L., Cluj-Napoca, 1994.

Miclea, Mircea, *Psihologie cognitivă. Metode teoretico-experimentale*, Editura Polirom, Iaşi, 1999.

Mihăilă, N., *Statistică matematică*, Editura Didactică şi Pedagogică, Bucureşti, 1965.

Mihoc, G.; Urseanu, V., *Sondaje şi estimaţii statistice*, Editura Tehnică, Bucureşti, 1977.

Mihoc, G.; Urseanu, V.; Ursianu, E., *Modele de analiză statistică*, Editura Ştiinţifică şi Enciclopedică, Bucureşti, 1982.

Moore, D.S.; McCabe, G.P., *Introduction to the Practice of Statistics*, (ediţia a III-a), W.H. Freeman and Company, New York, 1998.

Nowak, A., *Metode cantitative în psihologie şi sociologie*, Oscar Print, Bucureşti, 1998.

Pavelcu, Vasile, *Drama psihologiei*, Editura Didactică şi Pedagogică, Bucureşti, 1972.

Piaget, Jean, *Introduction à l'épistémologie génétique*, vol. II, Presses Universitaires de France, Paris, 1950.

Piaget, Jean, *Naşterea inteligenţei la copil*, Editura Didactică şi Pedagogică, Bucureşti, 1973.

Porojan, D., *Statistica şi teoria sondajelor*, Casa de editură şi presă „Şansa", Bucureşti, 1993.

Radu, I.; Miclea, M.; Moldovan, O.; Nemeş, S.; Szamoskozy, S., *Metodologia psihologică şi analiza datelor*, Editura Sincron, Cluj-Napoca, 1993.

Rancu, N.; Tövissi, L., *Statistică matematică cu aplicaţii în producţie*, Editura Academiei, Bucureşti, 1963.

Reber, A.S., *The Penguin Dictionary of Psychology*, Penguin Books, Londra, 1985.

Reuchlin, M., *Les méthodes quantitatives en psychologie*, Presses Universitaires de France, Paris, 1962.

Reuchlin, M., *Précis de statistique*, Presses Universitaires de France, Paris, 1976.

Robert, Michel (coord.), *Fondements et étapes de la recherche scientifique en psychologie*, Edisem, Maloine S.A., Paris, 1988.

Rotariu, T., *Metode statistice aplicate în ştiinţele sociale*, Editura Polirom, Iaşi, 1999.

Rouanet, Henry; Le Roux, Brigitte, *Statistique en sciences humaines. Exercices & Solutions*, Dunod, Paris, 1995.

Rude, Nathalie; Retel, O., *Statistique en psychologie*, In Press Éditions, Paris, 2000.

Stevens, Smith S., „Measurement", în Garry M. Maranell, *Scaling: A Sourcebook for Behavioral Scientist*, Aldine Publishing Company, Chicago, 1974.

Suppes, Patrick, *Metafizica probabilistă*, Editura Humanitas, Bucureşti, 1990.

Tiberghien, Guy, „Modèles de l'activité cognitive", în Caverni, Jean-Paul; Bastien, Claude; Mendelsohn, Patrick; Tiberghien, Guy, *Psychologie cognitive. Modèles et méthodes*, Presses Universitaires de Grenoble, Grenoble, 1991.

Vasilescu, I.P., *Statistică informatizată pentru ştiinţe despre om*, vol. 1, Editura Militară, Bucureşti, 1991.

Von Meyer, Harald, „Statistikangst und Einstellungen gegenüber Statistik bei Studierenden der Verhaltenwissenschaften", în *Forschungsforum*, Berichte aus Otto-Friedrich-Universität Bamberg *Mit Leib und Seele. Psychologie in Bamberg*, von Lothar, L. und Reinecker, H., Bamberg, 1994.

Vorberg, D., Blankenberger, S., „Alegerea testelor statistice şi măsurarea", în *Psihologia socială*, Buletinul Laboratorului „Psihologia câmpului social", 4, Universitatea „Al.I. Cuza", Editura Polirom, Iaşi, 1999.

Wottawa, H., *Psychologische Methodenlehre*, Juventa Verlag, München, 1977.

Zapan, Gheorghe, *Cunoaşterea şi aprecierea obiectivă a personalităţii*, Editura Ştiinţifică şi Enciclopedică, Bucureşti, 1984.

Collegium. Psihologie

au apărut:

Adrian Neculau (coord.) - *Psihologie socială. Aspecte contemporane*
Andrei Cosmovici - *Psihologie generală*
W. Doise, J.-C. Deschamps, G. Mugny - *Psihologie socială experimentală*
Adrian Neculau, Gilles Ferréol (coord.) - *Minoritari, marginali, excluşi*
Constantin Cucoş - *Minciună, contrafacere, simulare. O abordare psihopedagogică*
Mielu Zlate (coord.) - *Psihologia vieţii cotidiene*
R.Y. Bourhis, J.-P. Leyens (coord.) - *Stereotipuri, discriminare şi relaţii intergrupuri*
Serge Moscovici - *Psihologia socială sau maşina de fabricat zei*
Adrian Neculau (coord.) - *Câmpul universitar şi actorii săi*
Adrian Neculau (coord.) - *Psihologia câmpului social. Reprezentările sociale*
J. Barus-Michel, F. Giust-Desprairies, Luc Ridel - *Crize. Abordare psihosocială clinică*
Andrei Cosmovici, Luminiţa Iacob (coord.) - *Psihologie şcolară*
Şerban Ionescu - *Paisprezece abordări în psihopatologie*
Serge Moscovici (coord.) - *Psihologia socială a relaţiilor cu celălalt*
Willem Doise, Gabriel Mugny - *Psihologie socială şi dezvoltare cognitivă*
Ana Stoica-Constantin, Adrian Neculau (coord.) - *Psihosociologia rezolvării conflictului*
André Sirota - *Conduite perverse în grup*
Adrian Neculau, G. Ferréol (coord.) - *Psihosociologia schimbării*
Mircea Miclea - *Psihologie cognitivă. Modele teoretico-experimentale*
Mielu Zlate - *Psihologia mecanismelor cognitive*
Gilles Ferréol (coord.) - *Identitatea, cetăţenia şi legăturile sociale*
Mielu Zlate - *Introducere în psihologie*
Ion Dafinoiu - *Elemente de psihoterapie integrativă*
Eva Drozda-Senkowska - *Psihologia socială experimentală*
Cornel Havârneanu - *Cunoaşterea psihologică a persoanei. Posibilităţi de utilizare a computerului în psihologia aplicată*
Steve Duck - *Psihologia relaţiilor interumane. Teorie, concepte şi noţiuni fundamentale*
Valentin Clocotici, Aurel Stan - *Statistică aplicată în psihologie*

în pregătire:

Adrian Neculau, Pierre De Visscher (coord.) - *Dinamica grupurilor. Texte de bază*

Redactor: Adina Cobuz
Tehnoredactor: Paul Păduraru

Bun de tipar: septembrie 2000. Apărut 2001
Editura Polirom, B-dul Copou nr. 4 • P.O. Box 266, 6600, Iaşi
Tel. & Fax (032) 21.41.00; (032) 21.41.11;
(032) 21.74.40 (difuzare); E-mail: polirom@mail.dntis.ro
Bucureşti, B-dul I.C. Brătianu nr. 6, et. 7;
Tel.: (01) 313.89.78, E-mail: polirom@dnt.ro

Tiparul executat sub c-da 122
la S.C. GALINVEST S.R.L. GALAŢI
Str. Henri Coandă nr. 3
Tel.: 036/472381; 472927; Fax: 036/465471; 411862
E-mail: galinvest@xnet.ro
ROMÂNIA